Troubling thoughts

A sustained commitment to mental–health treatment for Fukushima evacuees could also help survivors of future disasters.

Natural disasters create bold images: winds blow, waves crash ashore, buildings burn. Yet some of the most important long-term consequences are invisible. Survivors can be gripped by anxiety, depression and post-traumatic stress disorder (PTSD) for years after a disaster — especially if they are unable to return home and rebuild.

This is exactly the situation faced by thousands of residents of eastern Fukushima prefecture in Japan, who fled their homes ahead of the meltdown of three reactors at the Fukushima Daiichi nuclear plant in March 2011. As described on page 290, many evacuees are now anxious and depressed about their long-term prospects. Researchers have found evidence that they suffer from elevated levels of anxiety and PTSD-related symptoms, and there are fears that substance abuse and other problems may be on the rise.

Mental health is a major component of the Fukushima Health Management Survey, the government effort to monitor survivors for adverse health effects from the accident. The survey includes a small but competent team of mental-health professionals who are chronicling mental-health and lifestyle issues. Last year, they heard from more than 90,000 evacuees who answered a brief questionnaire about their mental state. The results were not encouraging: levels of PTSD-like symptoms nearly a year after the accident were similar to those of workers who had responded to the 2001 attacks on the World Trade Center in New York.

The health survey's goal is not simply to record the effects of the accident, but also to support the evacuees. Counsellors targeted nearly 5,000 individuals for follow-up phone calls to discuss any mental-health problems. Unfortunately, a phone call may not be enough. The survey itself had less than a 50% response rate, and those who were called for follow-ups tended to stay on the line for just a few minutes — unwilling or unable to talk in detail about their problems.

The researchers involved in the health survey would like to do more. They want to conduct face-to-face interviews and set up counselling centres. But money is tight and the survey is already over budget. It currently spends about twice what the government has allocated, and budgetary wrangling between the prefectural and central governments could lead to further cuts. In this precarious financial environment, the team has found itself unable to hire long-term staff, or even print pamphlets about mental-health issues for the evacuees.

That is an unfortunate state of affairs because Fukushima presents a good opportunity to learn about the best ways to treat the mental-health problems of disaster victims. Although researchers have chronicled mental-health impacts from diverse crises ranging from 2005's Hurricane Katrina and the 2010 earthquake in Haiti to the ongoing Israeli–Palestinian conflict, the literature on how to actually treat these problems is fairly thin. Fukushima's evacuees are similar to survivors of other events in many ways. Nearly all suffer from sub-clinical symptoms that are difficult to diagnose. The size of the afflicted population makes one-on-one therapy impractical. And the

situation is complicated by a deep mistrust of the authorities, whom many evacuees blame for their present woes. But unlike many disaster zones, Fukushima is a developed region with a well-educated, well-documented and contactable population. Much could be learned by studying and treating these evacuees in the long term.

And the evacuees will need that help. Unlike the survivors of many other disasters (including those who were affected by the tsunami that sparked the nuclear meltdown), Fukushima's nuclear evacuees live in fear of radiation exposure. They worry that they or their children may fall ill from the accident, and their anxiety could grow as the years pass. Nearly half of all Japanese people will develop cancer at some point in their lives, and evacuees will wonder whether that cancer is connected to the accident. Studies of mothers who were evacuated from the 1986 nuclear disaster in Chernobyl, Ukraine, show that this anxiety for their children can last for decades, and may never entirely go away.

> *"Although researchers have chronicled mental–health impacts from diverse crises, the literature on how to treat these problems is thin."*

Given all this, the Fukushima Health Management Survey deserves continued and even increased support from the government. Survey scientists should also seek lasting collaborations with researchers outside Japan so that the lessons learned there, so painfully earned, can be shared with the world. ∎

Knowledge trades

Institutions must carefully evaluate their researchers' relationships with Wall Street.

Few researchers would care to put a price on their professional reputation. Yet when neurologist Sidney Gilman decided to divulge confidential clinical-trial data to a hedge-fund manager, he did exactly that. Court documents show that Gilman earned more than US$100,000 for his illegal tips about the failure of an experimental drug for Alzheimer's disease; the hedge fund made $276 million in direct gains and avoided losses. After he was caught, Gilman lost not only the cash but also his career, retiring from his position at the University of Michigan Medical School in Ann Arbor.

In recent conversations with a *Nature* reporter, several academics converged in their assessment of Gilman's case, which became the biggest insider-trading case in US history: "stupid".

Nevertheless, many of these same academics continue to serve as

consultants for the investment industry, where knowledge is money and confidential knowledge is the most precious currency of all. Like Gilman, they participate in 'expert network' firms that bring together academic specialists and clients who seek technical information (see page 280). Such firms do not work exclusively for the investment industry, but hedge funds make up a sizeable part of their business. Expert networks make it easier for researchers to dabble at advising Wall Street, often for clients who do not disclose the companies for which they work.

Undoubtedly, the vast majority of consultations for financial firms do not result in illegal exchanges. Gilman and his alleged hedge-fund co-conspirator were clearly mindful of their transgressions and went to some lengths to evade the barriers to insider trading erected by the expert network that united them in the first place. Experts contacted by *Nature* were confident that they had never divulged confidential information. They were less confident about others, noting that less experienced or over-eager colleagues might be prone to a slip of the tongue here and there.

A small slip can move markets. Telltale cues from body language (a shift in a chair) and tone of voice (a hesitation, a cough) can speak volumes, as can the unguarded answer to one acute question slipped into an otherwise innocuous conversation. Indeed, journalists might find the techniques used by hedge-fund managers to dig out confidential information uncomfortably familiar. Many academics are trained to handle the press; few receive education in how to deal with the financial industry.

The cavalier attitude towards this work is disturbing. For busy physicians, recruitment letters from expert networks are part of a steady flow of surveys and consulting requests that clog their inboxes. Some of the researchers interviewed by *Nature* could not remember which expert networking firms they consult for. Consultations were often viewed as an easy way to pick up a little extra cash when time permits, and, if lucky, perhaps have an engaging conversation along the way.

This casual approach extends to institutions, many of which are well versed in negotiating the rocky road of conflicts of interest raised by consultations for the drug or medical-device industries, but have not explored the issues raised when advising hedge funds. Universities are already wrestling with the mounting requirements from federal funders for the reporting of potential conflicts of interest, and are unlikely to welcome yet another category of extracurricular activities to monitor.

All involved must take the relationship between researchers and Wall Street more seriously. Institutions should discuss the risks involved and, when warranted, take a proactive stance, perhaps using the Cleveland Clinic in Ohio as a guide. Since 2005, the clinic has instituted a special level of legal review for relationships between faculty members and the investment industry. Physicians who embark on these relationships are given special educational material — which might soon include the newspaper accounts of the Gilman case.

> *"Avoidance may be the best strategy to prevent accidental leaks."*

As US regulators cracked down on insider trading, some researchers cut their ties with the financial industry and expert networks for fear of being tainted by association. Those who still consult for Wall Street often say that they do so to help guide investment in their field. That aim is laudable, but it cannot be used to justify consultation on topics that could overlap with privileged information. For example, a researcher who works on a clinical trial for an experimental diabetes drug should think twice before consulting with a hedge fund about diabetes drugs.

Avoidance may be the best strategy to prevent accidental leaks. The Gerson Lehrman Group, the expert network based in New York that employed Gilman, has policies in place to protect consultants from the consequences of unintended disclosure. Two stand out as potentially the most powerful: an academic can refuse a consultation if the subject matter might tread near confidential information; and he or she can abort a consultation — and still get paid — if a client presses for insider knowledge. So the solution is simple: when asked to disclose confidential information, hang up the phone. ∎

Natural history

Age-old field methods can tell us more about animal behaviour than can laboratory models.

In the autumn of 1927, the biologist Francis Sumner spent two months in Florida and Alabama trapping wild oldfield mice for studies of skin pigmentation. With the advice of local farmers, Sumner managed to shoo hundreds of rodents out of their burrows.

Sumner also took the opportunity to document the complexity of the vacated mouseholes, detailed in an article published two years later (Sumner, F. B. & Karol, J. J. *J. Mammal.* **10**, 213–215; 1929). The creatures' burrows included a long entrance tunnel leading to a nest and, in case of an invading snake, a secondary escape tunnel that didn't quite reach the surface, a metre or so from the entrance. Other ethologists have since characterized the humble mouse burrow, the structure of which is seen as a model of complex animal behaviour.

Fast-forward almost a century, and a team led by Hopi Hoekstra of Harvard University in Cambridge, Massachusetts, has elegantly unpicked the genetic basis of this behaviour using a cross-breeding design and cutting-edge genotyping methods. Such work, published in this issue (see pages 284 and 402), should appeal to more than just mouse fanciers. Like few papers before, the work shows how long-forgotten field observations, evolutionary theory and molecular genetics can all be brought to bear on a single question.

We have learned much about the physiology of behaviour from model organisms such as laboratory mice — for example, the discovery of genes that determine circadian rhythms, which revealed important mechanisms underlying behaviours such as sleep. But decades of selection for convenient traits such as docility have made laboratory models less than ideal for studying the evolution of complex behaviours. They tell biologists little about the vast behavioural differences that can exist between closely related animals, probably as a result of natural and sexual selection.

Scientists interested in probing the behaviour of wild animals can follow Hoekstra's lead and pick animals and behaviours with a rich history of observation and striking differences between close relatives. Decades-old observations of ant behaviour, including those by Edward O. Wilson, culminated in the discovery, published online in *Nature* today, that a social chromosome explains why some red imported fire ant colonies have one queen, whereas others accept multiple queens.

There are, of course, risks to tackling behavioural genomics in wild animals. The ultimate proof of any gene's role in a specific behaviour involves knocking in or out the gene to remove or endow that behaviour. Such experiments are a challenge even in model organisms, and so far few precedents have been set in non-model species.

Model organisms, imperfect as they are when it comes to studying some behaviours, have focused attention on a handful of organisms. If every interesting animal becomes fair game, there is a risk that behavioural genetics will be fragmented. "If everyone does it in their own species, it will not be a very productive type of enterprise," says Laurent Keller, a geneticist at the University of Lausanne, Switzerland, who led the ant research. He suggests that scientists converge on a set of wild animals in which to intensely study behaviour. If the latest research is any indication, such animals will be no strangers to historians of biology. ∎

↻ NATURE.COM
To comment online, click on Editorials at:
go.nature.com/xhunqv

China has the capacity to lead in carbon trading

Pilot schemes launched this year could be the start of a world-class system — if the country can solve its data-gathering problems, says **Qiang Wang**.

The value of the world's carbon market fell for the first time last year. More than one-third was wiped from the price of carbon credits in a plunge that reflects the continued global economic crisis and uncertainty over the future of emissions-trading schemes.

Is it a good time to buy carbon credits? Perhaps not yet, but some shoots of recovery are visible, not least in my own country of China. Largely unnoticed in the West, Chinese carbon trading is getting up and running. In just two years, officials have designed and started to implement seven trading trials that cover around one-third of China's gross domestic product and one-fifth of its energy use. If successful, the schemes will show that emissions trading could be a powerful way for China to control its greenhouse-gas emissions.

The first trades in one of these schemes took place in September 2012, when four cement-manufacturing companies in the southern, industrial region of Guangdong province invested several million dollars each in carbon-pollution permits, which they will need to expand operations. The Guangdong scheme is expected to cover more than 800 companies that each emit more than 20,000 tonnes of carbon dioxide a year across nine industries, including the energy-intensive steel and power sectors. These firms account for more than 40% of the power used in the province. The Guangdong carbon market alone will regulate some 277 million tonnes of CO_2 emissions by 2015, almost equal to Ukraine's total annual CO_2 emissions.

China plans to open six further regional emissions-trading schemes this year, in the province of Hubei and in the municipalities of Beijing, Tianjin, Shanghai, Chongqing and Shenzhen. It plans to expand and link them until they form a nationwide scheme by the end of the decade; that would then link to international markets.

China is not alone. Australia and South Korea are scheduled to open their own carbon markets in 2015, and California launched its first carbon-allowances auction in November 2012.

China may not seem a natural home for carbon trading. With heavy government intervention, significant state ownership of enterprises and a culture of distrust in business, the country remains far from a true market economy. Until now, China's experience of carbon trading has been almost exclusively under the Clean Development Mechanism, the international carbon-trading scheme set up under the Kyoto Protocol, to which it is the leading supplier of credits. Only a few symbolic transactions of carbon offsets have been made since the Beijing, Tianjin and Shanghai voluntary emissions-trading exchanges opened in 2008.

However, China's political system could let a carbon market grow faster than anywhere else. Once Chinese leaders have accepted a concept, opposition is steamrollered and changes are implemented much more quickly and broadly than is possible in societies in which policy-making is based on a balance of the interests from different stakeholders. Stable policies are crucial to the success of emissions trading, which — unlike most markets — grows from the top down.

Challenges lie ahead for emissions trading in China. The country needs to develop and enforce proper legislation and regulations to measure, report and verify carbon emissions from industrial sites. It needs to build an effective and accountable framework to oversee the reporting and trading of carbon credits. Most urgently, China needs to look at how it collects and analyses data on carbon emissions.

The credibility of China's statistics on energy use and carbon emissions has long been questioned, partly because numbers calculated using top-down and bottom-up statistics do not match. The discrepancies can be very large. For example, last year, scientists compared Chinese CO_2 emissions as calculated using top-down energy data from the National Bureau of Statistics of China with those calculated using bottom-up data from the 30 provincial statistics bureaus. For 2010, the bottom-up figure was larger by an amount equivalent to the total annual CO_2 emissions of Japan, about 5% of the global total (D. Guan *et al. Nature Clim. Change* **2,** 672–675; 2012). Without accurate numbers, the first deal of the Guangdong trading scheme was based on expected future carbon emissions, rather than historical data.

China should set up a reliable system for gathering carbon data as soon as possible, and build a comprehensive database of carbon emissions. Standard international methods for carbon trading to reduce emissions demonstrate how to set up a baseline, show that any emissions change is genuine and calculate the reductions.

Reliable data-gathering will require effort from the scientific community at home and abroad, to improve statistical methodology. It will also need political action. The organizations that collate and publish carbon statistics should be made independent of possible government interference. And China needs specific laws to ensure honest reporting and transparency, and to punish those who make fraudulent or misleading claims about their carbon emissions.

If China can build a workable and credible emissions-trading scheme, it will benefit not just one country, but the entire world. ■

> ## CHINA'S POLITICAL SYSTEM COULD LET A CARBON MARKET GROW FASTER THAN ANYWHERE ELSE.

Qiang Wang *is director of the Western Research Center for Energy and Eco-Environmental Policy at the Xinjiang Institute of Ecology and Geography of the Chinese Academy of Sciences in Urumqi.*
e-mail: wang_q@ms.xjb.ac.cn

RESEARCH HIGHLIGHTS

Selections from the scientific literature

Machine mimics ribosome

The cell's protein-building complex has inspired the development of a molecular machine that links amino acids together.

David Leigh at the University of Manchester, UK, and his colleagues based their 'nanomachine' on the ribosome, a complex of proteins and RNA that translates the genetic code into proteins. But, compared with the ribosome, their machine is primitive and slow, and can assemble only very short protein chains called peptides. The nanomachine is based on a rotaxane, which is a large molecular ring threaded onto a molecule that acts as an axle. In the experiment, the axle is lined with three amino acids, and a chain of three amino acids hangs from the outer edge of the ring. When the machine is heated, an amino acid from the axle is transferred to the end of the hanging chain. The ring can then move along the axle, repeating the reaction.

Running about 10^{18} molecular machines at once produces tens of milligrams of peptide.
Science 339, 189–193 (2013)
For a longer story on this research, see go.nature.com/p3rtai

---

Migration from India to Australia

A genetic analysis indicates that humans migrated from India to Australia around 4,000 years ago. This contradicts the prevailing view that, after its initial colonization, Australia had little contact with the rest of the world before the late eighteenth century.

Irina Pugach at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, and her colleagues analysed genome-wide genetic variation from 344 people, including Australian Aboriginals and individuals from Africa, India and island Southeast Asia. The authors found evidence of a migration out of Africa into Australia at least 36,000 years ago, and one from India about 4,230 years ago.

The Australian archaeological record shows rapid changes in tool technology around 4,000 years ago, leading the authors to suggest that the migrants could have brought in the new technology.
Proc. Natl Acad. Sci. USA http://dx.doi.org/10.1073/pnas.1211927110 (2013)

Cicadas emerge when predators decline

Periodical cicadas, which spend most of their lives underground and occasionally emerge en masse, may time their appearances with periodic dips in predator numbers.

Why several species of the insect (*Magicicada* spp; pictured) surge to the surface of North American forests after 13 or 17 years underground has been a mystery. Walt Koenig at Cornell University in Ithaca, New York, and Andrew Liebhold of the US Department of Agriculture Forest Service in Morgantown, West Virginia, analysed data on the population sizes of 15 bird species between 1966 and 2010 to estimate the predation pressure on the insects. They conclude that the cicadas somehow set their predators' populations into cycles that reduce bird numbers during their next foray above ground.

This could be explained by the insects' long-lasting effect on forest ecology influencing factors such as nutrient levels and tree growth for several years after their emergence, the authors suggest.
Am. Nat. 181, 145–149 (2013)

Hearing restored with new hair cells

An experimental drug can restore hearing in deaf mice by regenerating sound-sensitive cells in the inner ear.

Excessive noise permanently damages hair cells that conduct sound to the brain and are found in an inner-ear structure called the cochlea. A team led by Albert Edge at Harvard Medical School in Boston, Massachusetts, has identified a molecule that can convert other cochlear cells into hair cells. The compound, named LY411575, blocks a biochemical pathway called Notch, which normally prevents supporting cells in the cochlea from developing

Machine mimics ribosome

The cell's protein-building complex has inspired the development of a molecular machine that links amino acids together.

David Leigh at the University of Manchester, UK, and his colleagues based their 'nanomachine' on the ribosome, a complex of proteins and RNA that translates the genetic code into proteins. But, compared with the ribosome, their machine is primitive and slow, and can assemble only very short protein chains called peptides. The nanomachine is based on a rotaxane, which is a large molecular ring threaded onto a molecule that acts as an axle. In the experiment, the axle is lined with three amino acids, and a chain of three amino acids hangs from the outer edge of the ring. When the machine is heated, an amino acid from the axle is transferred to the end of the hanging chain. The ring can then move along the axle, repeating the reaction.

Running about 10^{18} molecular machines at once produces tens of milligrams of peptide.
Science **339**, 189–193 (2013)
For a longer story on this research, see go.nature.com/p3rtai

Migration from India to Australia

A genetic analysis indicates that humans migrated from India to Australia around 4,000 years ago. This contradicts the prevailing view that, after its initial colonization, Australia had little contact with the rest of the world before the late

ED RESCHKE/PETER ARNOLD/GETTY

Cicadas emerge when predators decline

Periodical cicadas, which spend most of their lives underground and occasionally emerge en masse, may time their appearances with periodic dips in predator numbers.

Why several species of the insect (*Magicicada* spp; pictured) surge to the surface of North American forests after 13 or 17 years underground has been a mystery. Walt Koenig at Cornell University in Ithaca, New York, and Andrew Liebhold of the US Department of Agriculture Forest Service in Morgantown, West Virginia, analysed data on the population

sizes of 15 bird species between 1966 and 2010 to estimate the predation pressure on the insects. They conclude that the cicadas somehow set their predators' populations into cycles that reduce bird numbers during their next foray above ground.

This could be explained by the insects' long-lasting effect on forest ecology influencing factors such as nutrient levels and tree growth for several years after their emergence, the authors suggest.
Am. Nat. **181**, 145–149 (2013)

eighteenth century.

Irina Pugach at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, and her colleagues analysed genome-wide genetic variation from 344 people, including Australian Aboriginals and individuals from Africa, India and island Southeast Asia. The authors found evidence of a migration out of Africa into Australia at least 36,000 years ago, and one from India about 4,230 years ago.

The Australian archaeological record

shows rapid changes in tool technology around 4,000 years ago, leading the authors to suggest that the migrants could have brought in the new technology.
Proc. Natl Acad. Sci. USA
http://dx.doi.org/10.1073/pnas.1211927110 (2013)

Hearing restored with new hair cells

An experimental drug can restore hearing in deaf mice by regenerating sound-sensitive

cells in the inner ear.

Excessive noise permanently damages hair cells that conduct sound to the brain and are found in an inner-ear structure called the cochlea. A team led by Albert Edge at Harvard Medical School in Boston, Massachusetts, has identified a molecule that can convert other cochlear cells into hair cells. The compound, named LY411575, blocks a biochemical pathway called Notch, which normally prevents supporting cells in the cochlea from developing

into hair cells. When applied to the ears of mice deafened by loud sounds, the molecule triggered the generation of new hair cells. By three months after this treatment, the rodents' hearing had been partly restored.

LY411575 would probably need to be delivered locally to the ear in humans to avoid side effects, the authors say.
Neuron 77, **58–69 (2013)**

Mothers call for parenting help

Male laboratory mice do not spontaneously care for their offspring, but they can exhibit some paternal behaviour in response to signals from their mates.

Haruhiro Higashida at Kanazawa University in Japan and his colleagues found that when female mice were separated from their young, they could encourage their male partners to adopt parental behaviour — such as carrying wandering pups back to the nest (**pictured**) — using high-pitched vocalizations and pheromones.

About 60% of the males tested retrieved pups after hearing a recording of 38-kilohertz distress calls of females removed from their pups. The team found similar results when males were placed in cages that had previously held females separated from their young. Paternal behaviour disappeared when the researchers simultaneously blocked hearing and olfaction in the males.
Nature Commun. 4, **1346 (2013)**

ARCO IMAGES GMBH/ALAMY

Spinning strong, conductive carbon

Fibres spun from carbon nanotubes can be as electrically conductive as metal wires, yet as strong as conventional carbon fibres.

Individual carbon nanotubes are strong, stiff and exceptionally conductive — but spinning them into larger filaments adds defects, impurities and misalignments that compromise the fibres' physical properties. Matteo Pasquali at Rice University in Houston, Texas, and his colleagues have improved the spinning process by adapting techniques used to make industrial fibres such as Kevlar. The authors extruded filaments from nanotubes dissolved in an acid and wound them into fibres of dense, well-aligned tubes that are stronger than copper and almost as conductive.

The fibres could be manufactured on a large scale and have a variety of applications in electronics, the authors say.
Science 339, **182–186 (2013)**

Impede a protein to reprogram

Blockade of just one protein is sufficient to change skin cells into neurons.

Several research groups have already converted one cell type to another by adding, for example, specific genes. Now researchers led by Xiang-Dong Fu at the University of California, San Diego, have found another approach. They used small RNA molecules to repress an RNA-binding protein called PTB. Decreased levels of this protein led to the activation of genes that can convert skin cells to neurons. PTB influences the stability of RNA molecules and coordinates molecules called microRNAs that regulate gene expression.

The approach is a

Warming matches predictions

★ HIGHLY READ on iopscience. iop.org in December

Temperature data collected since 1990 are in good agreement with model-based estimates of global warming made by the Intergovernmental Panel on Climate Change (IPCC).

Stefan Rahmstorf of the Potsdam Institute for Climate Impact Research in Germany and his colleagues removed short-term temperature fluctuations or 'noise', which result from events such as volcanic eruptions, from a 1980–2011 time-series of annual global temperatures. The adjusted warming trend of 0.16 °C per decade closely matches the projections made by the IPCC in its reports in 2001 and 2007.

The global sea level, however, has since 1993 been rising 60% faster than anticipated. Future sea-level rise could exceed even the highest value — some 60 centimetres by the end of the century — projected by the IPCC in its 2007 report, the team cautions.
Environ. Res. Lett . 7, **044035 (2012)**

GUDRUN HERZNER/PROC. NATL ACAD. SCI.

convenient way to reprogram other cell types to neurons, and provides a glimpse of the complex mechanisms involved in the control of gene expression.
Cell http://dx.doi.org/10.1016/j. cell.2012.11.045 (2013)

Wasp parasites keep hosts clean

By peering into the bodies of American cockroaches, researchers have found that a parasitic larva disinfects its host with antibacterial secretions.

Larvae of the emerald cockroach wasp (*Ampulex compressa*; **pictured**) feed on the innards of the American cockroach (*Periplaneta americana*), but also have to contend with bacteria that live in their host's tissues. Gudrun Herzner at the University

of Regensburg in Germany and her colleagues installed transparent panels into the sides of parasitized roaches and observed that the wasp larvae secrete large amounts of a clear liquid from their mouthparts onto their hosts' tissues.

Analysis of the liquid revealed the presence of the chemicals mellein and micromolide. These substances slow the growth of certain microorganisms, including the bacterium *Serratia marcescens*, which can kill insect larvae and was also isolated from the cockroaches.
Proc. Natl Acad. Sci. USA http://dx.doi.org/10.1073/ pnas.1213384110 (2013) **For a longer story on this research, see go.nature.com/vxttfo**

↻ **NATURE.COM**
For the latest research published by *Nature* visit:
www.nature.com/latestresearch

into hair cells. When applied to the ears of mice deafened by loud sounds, the molecule triggered the generation of new hair cells. By three months after this treatment, the rodents' hearing had been partly restored.

LY411575 would probably need to be delivered locally to the ear in humans to avoid side effects, the authors say.
Neuron 77, **58–69** (2013)

ANIMAL BEHAVIOUR

Mothers call for parenting help

Male laboratory mice do not spontaneously care for their offspring, but they can exhibit some paternal behaviour in response to signals from their mates.

Haruhiro Higashida at Kanazawa University in Japan and his colleagues found that when female mice were separated from their young, they could encourage their male partners to adopt parental behaviour — such as carrying wandering pups back to the nest (**pictured**) — using high-pitched vocalizations and pheromones.

About 60% of the males tested retrieved pups after hearing a recording of 38-kilohertz distress calls of females removed from their pups. The team found similar results when males were placed in cages that had previously held females separated from their young. Paternal behaviour disappeared when the researchers simultaneously blocked hearing and olfaction in the males.
Nature Commun. 4, **1346** (2013)

ARCO IMAGES GMBH/ALAMY

MATERIALS

Spinning strong, conductive carbon

Fibres spun from carbon nanotubes can be as electrically conductive as metal wires, yet as strong as conventional carbon fibres.

Individual carbon nanotubes are strong, stiff and exceptionally conductive — but spinning them into larger filaments adds defects, impurities and misalignments that compromise the fibres' physical properties. Matteo Pasquali at Rice University in Houston, Texas, and his colleagues have improved the spinning process by adapting techniques used to make industrial fibres such as Kevlar. The authors extruded filaments from nanotubes dissolved in an acid and wound them into fibres of dense, well-aligned tubes that are stronger than copper and almost as conductive.

The fibres could be manufactured on a large scale and have a variety of applications in electronics, the authors say.
Science 339, **182–186** (2013)

CELL FATE

Impede a protein to reprogram

Blockade of just one protein is sufficient to change skin cells into neurons.

Several research groups have already converted one cell type to another by adding, for example, specific genes. Now researchers led by Xiang-Dong Fu at the University of California, San Diego, have found another approach. They used small RNA molecules to repress an RNA-binding protein called PTB. Decreased levels of this protein led to the activation of genes that can convert skin cells to neurons. PTB influences the stability of RNA molecules and coordinates molecules called microRNAs that regulate gene expression.

The approach is a convenient way to reprogram other cell types to neurons, and provides a glimpse of the complex mechanisms involved in the control of gene expression.
Cell http://dx.doi.org/10.1016/j.cell.2012.11.045 (2013)

ANIMAL BEHAVIOUR

Wasp parasites keep hosts clean

By peering into the bodies of American cockroaches, researchers have found that a parasitic larva disinfects its host with antibacterial secretions.

Larvae of the emerald cockroach wasp (*Ampulex compressa*; **pictured**) feed on the innards of the American cockroach (*Periplaneta americana*), but also have to contend with bacteria that live in their host's tissues. Gudrun Herzner at the University of Regensburg in Germany and her colleagues installed transparent panels into the sides of parasitized roaches and observed that the wasp larvae secrete large amounts of a clear liquid from their mouthparts onto their hosts' tissues.

Analysis of the liquid revealed the presence of the chemicals mellein and micromolide. These substances slow the growth of certain microorganisms, including the bacterium *Serratia marcescens*, which can kill insect larvae and was also isolated from the cockroaches.
Proc. Natl Acad. Sci. USA http://dx.doi.org/10.1073/pnas.1213384110 (2013)
For a longer story on this research, see go.nature.com/vxttfo

GUDRUN HERZNER/PROC. NATL ACAD. SCI.

CLIMATE SCIENCE

Warming matches predictions

⭐ **HIGHLY READ** on iopscience.iop.org in December

Temperature data collected since 1990 are in good agreement with model-based estimates of global warming made by the Intergovernmental Panel on Climate Change (IPCC).

Stefan Rahmstorf of the Potsdam Institute for Climate Impact Research in Germany and his colleagues removed short-term temperature fluctuations or 'noise', which result from events such as volcanic eruptions, from a 1980–2011 time-series of annual global temperatures. The adjusted warming trend of 0.16 °C per decade closely matches the projections made by the IPCC in its reports in 2001 and 2007.

The global sea level, however, has since 1993 been rising 60% faster than anticipated. Future sea-level rise could exceed even the highest value — some 60 centimetres by the end of the century — projected by the IPCC in its 2007 report, the team cautions.
Environ. Res. Lett . 7, **044035** (2012)

⭮ **NATURE.COM**
For the latest research published by *Nature* visit:
www.nature.com/latestresearch

into hair cells. When applied to the ears of mice deafened by loud sounds, the molecule triggered the generation of new hair cells. By three months after this treatment, the rodents' hearing had been partly restored.

LY411575 would probably need to be delivered locally to the ear in humans to avoid side effects, the authors say.
Neuron 77, **58–69 (2013)**

Mothers call for parenting help

Male laboratory mice do not spontaneously care for their offspring, but they can exhibit some paternal behaviour in response to signals from their mates.

Haruhiro Higashida at Kanazawa University in Japan and his colleagues found that when female mice were separated from their young, they could encourage their male partners to adopt parental behaviour — such as carrying wandering pups back to the nest (**pictured**) — using high-pitched vocalizations and pheromones.

About 60% of the males tested retrieved pups after hearing a recording of 38-kilohertz distress calls of females removed from their pups. The team found similar results when males were placed in cages that had previously held females separated from their young. Paternal behaviour disappeared when the researchers simultaneously blocked hearing and olfaction in the males.
Nature Commun. 4, **1346 (2013)**

ARCO IMAGES GMBH/ALAMY

MATERIALS

Spinning strong, conductive carbon

Fibres spun from carbon nanotubes can be as electrically conductive as metal wires, yet as strong as conventional carbon fibres.

Individual carbon nanotubes are strong, stiff and exceptionally conductive — but spinning them into larger filaments adds defects, impurities and misalignments that compromise the fibres' physical properties. Matteo Pasquali at Rice University in Houston, Texas, and his colleagues have improved the spinning process by adapting techniques used to make industrial fibres such as Kevlar. The authors extruded filaments from nanotubes dissolved in an acid and wound them into fibres of dense, well-aligned tubes that are stronger than copper and almost as conductive.

The fibres could be manufactured on a large scale and have a variety of applications in electronics, the authors say.
Science 339, **182–186 (2013)**

CELL FATE

Impede a protein to reprogram

Blockade of just one protein is sufficient to change skin cells into neurons.

Several research groups have already converted one cell type to another by adding, for example, specific genes. Now researchers led by Xiang-Dong Fu at the University of California, San Diego, have found another approach. They used small RNA molecules to repress an RNA-binding protein called PTB. Decreased levels of this protein led to the activation of genes that can convert skin cells to neurons. PTB influences the stability of RNA molecules and coordinates molecules called microRNAs that regulate gene expression.

The approach is a convenient way to reprogram other cell types to neurons, and provides a glimpse of the complex mechanisms involved in the control of gene expression.
Cell http://dx.doi.org/10.1016/j. cell.2012.11.045 (2013)

Wasp parasites keep hosts clean

By peering into the bodies of American cockroaches, researchers have found that a parasitic larva disinfects its host with antibacterial secretions.

Larvae of the emerald cockroach wasp (*Ampulex compressa*; **pictured**) feed on the innards of the American cockroach (*Periplaneta americana*), but also have to contend with bacteria that live in their host's tissues. Gudrun Herzner at the University of Regensburg in Germany and her colleagues installed transparent panels into the sides of parasitized roaches and observed that the wasp larvae secrete large amounts of a clear liquid from their mouthparts onto their hosts' tissues.

Analysis of the liquid revealed the presence of the chemicals mellein and micromolide. These substances slow the growth of certain microorganisms, including the bacterium *Serratia marcescens*, which can kill insect larvae and was also isolated from the cockroaches.
Proc. Natl Acad. Sci. USA http://dx.doi.org/10.1073/ pnas.1213384110 (2013)
For a longer story on this research, see go.nature.com/vxttfo

↻ **NATURE.COM**
For the latest research published by *Nature* visit:
www.nature.com/latestresearch

CLIMATE SCIENCE

Warming matches predictions

★ **HIGHLY READ** on iopscience. iop.org in December

Temperature data collected since 1990 are in good agreement with model-based estimates of global warming made by the Intergovernmental Panel on Climate Change (IPCC).

Stefan Rahmstorf of the Potsdam Institute for Climate Impact Research in Germany and his colleagues removed short-term temperature fluctuations or 'noise', which result from events such as volcanic eruptions, from a 1980–2011 time-series of annual global temperatures. The adjusted warming trend of 0.16 °C per decade closely matches the projections made by the IPCC in its reports in 2001 and 2007.

The global sea level, however, has since 1993 been rising 60% faster than anticipated. Future sea-level rise could exceed even the highest value — some 60 centimetres by the end of the century — projected by the IPCC in its 2007 report, the team cautions.
Environ. Res. Lett. 7, **044035 (2012)**

GUDRUN HERZNER/PROC. NATL ACAD. SCI.

RESEARCH

Antarctic lakes

A team of US-led Antarctic explorers was preparing to drill through to the subglacial Lake Whillans as *Nature* went to press this week. The pristine lake is buried beneath some 800 metres of ice at the margin of the West Antarctic Ice Sheet. Meanwhile, Russian scientists who last February reached Antarctica's largest subglacial lake, Lake Vostok, announced on 10 January that they had recovered further samples of lake water, which had moved up the borehole and refrozen onto the drill bit. See go.nature.com/ut2hxw for more.

US climate audit

A US climate-assessment advisory committee released its draft report on 11 January, documenting a range of global-warming impacts across the country and making projections. Issued under the auspices of the inter-agency Global Change Research Program, the document is available for public comment until 12 April, and will serve as the basis for the United States' third national climate assessment, which is expected to be released late this year or early in 2014. See go.nature.com/k9p4lm for more.

EVENTS

Bush-fire damage

A fierce bush fire threatened Australia's Siding Spring Observatory in New South Wales last week, but initial assessments indicated that no telescopes were seriously damaged, the Australian National University (which runs the site) said on 14 January. Some buildings were severely affected at the observatory, which will be

Record smog blankets Beijing

A year after Beijing's municipal government said that it would report more realistic assessments of the city's air quality, pollution in China's capital soared to unprecedented heights. City officials last week asked construction sites and industries to limit their activities and told residents to stay indoors, as concentrations of fine particles (those less than 2.5 micrometres across) rose to record levels many times above what the World Health Organization considers 'hazardous'. With smog blanketing cities across northern China, state media urged the government to take action on air pollution.

closed for two weeks. The fire was one of many across Australia, where long-standing temperature records were broken last week during a summer heatwave.

Warmest year

The contiguous United States chalked up its warmest year on record last year, blowing past the previous record set in 1998. The average temperature, 12.9 °C, was 1.8 °C above the average for the twentieth century and 0.56 °C above the previous record. Global data are not yet available, but the latest analysis by the US National Oceanic and Atmospheric Administration suggests that, in 2012, January to November was the eighth warmest period on record.

POLICY

GM transparency

The European Food Safety Authority announced on 14 January that it was starting a broad transparency initiative, designed to make data from its risk assessments more widely available. The agency also released the first batch of data: a file containing documents and data submitted by biotechnology giant Monsanto for the 2003 authorization of its genetically modified maize (corn) NK603. See go.nature. com/e7sgdd for more.

Primates grounded

United Airlines, the world's largest carrier, will no longer ship non-human primates to research labs. The airline clarified its policy on 8 January, following pressure from activists at People for the Ethical Treatment of Animals. With the adoption of similar rules by Air Canada last month, only four major airlines now say they will fly research primates: Air France, China Eastern Airlines, Philippine Airlines and Vietnam Airlines. See go.nature.com/agkgbl for more.

Japanese windfall

Japanese scientists may benefit from the ¥10.3-trillion (US$120-billion) supplementary budget proposed by the Liberal Democratic Party, which regained power in December and seems set on reversing some of the austerity that has characterized Japan's leadership over the past few years. Of

¥570 billion requested by the science and education ministry, universities will receive ¥180 billion for technology transfer, and research on induced pluripotent stem cells could receive ¥20 billion. The budget is expected to be passed next month. See go.nature.com/7lhmxg for more.

California increase

California's public universities, which have suffered years of funding cuts, were awarded a US$500-million boost in the state's proposed 2013–14 budget, announced on 10 January. The generosity comes after voters last year approved a temporary tax increase to stave off deep cuts to public education. Despite the rise, general state funds stand more than $1.2 billion lower than in 2007–08, a shortfall that universities have made up for by increasing student tuition fees (see *Nature* **480**, 164; 2011). Both Democrats and Republicans expressed support for the budget, which the state legislature must approve before it comes into effect in July.

Red–Dead link

A proposal to construct a 180-kilometre-long conduit from the Red Sea to the Dead Sea is economically and environmentally feasible, according to draft studies posted online last week. Israel, Jordan and the Palestinian

Authority had asked the World Bank to look at the plan in 2005, with the idea that channelling water downhill to restore the salty and rapidly shrinking Dead Sea would also generate hydroelectricity to run a desalination plant (see *Nature* **464**, 1118–1120; 2010). The World Bank recommended a pipeline as the best option, at a cost of around US$10 billion. See go.nature.com/if8uyu for more.

US energy hub

The US Department of Energy has launched its fifth Energy Innovation Hub, a research centre focusing on rare-earth minerals and materials that are crucial to the energy system, including advanced batteries, wind turbines and solar panels. Based at the department's Ames Laboratory in Iowa, the Critical Materials Institute will receive up to US$120 million over five years. The institute, announced on 9 January, follows earlier hubs focusing on advanced nuclear reactors, energy storage, energy-efficient buildings and the development of fuels from sunlight. The hub model was the brainchild of energy secretary Steven Chu.

Open-access tribute

A campaign on Twitter is calling for researchers to post their papers openly online,

at hashtag #pdftribute, in commemoration of Aaron Swartz, a 26-year-old programmer and Internet activist who committed suicide on 11 January. Swartz (**pictured**), who pushed for open access to court records, scholarly material and other web information, was facing criminal charges for downloading some 4 million articles from the not-for-profit scholarly archive JSTOR. Academics have joined a wave of indignation at the charges, which many argue contributed to his suicide. See go.nature.com/5wmeld for more.

Energy spending

Global investment in clean energy fell by 11% to US$268.7 billion in 2012, according to figures from analysts Bloomberg New Energy Finance on 14 January. Around $30 billion of investment went to research

20–25 JANUARY
In Tromsø, Norway, politicians and researchers discuss the geopolitics and changing ecosystems of the Arctic, in the 'Arctic Frontiers' conference.
go.nature.com/kjorrb

24 JANUARY
The Natural History Museum in London launches its year-long celebration of the work of Alfred Russel Wallace — the co-discoverer of evolution by natural selection. This year is the centenary of his death.
go.nature.com/dxbt9r

and development, with the rest mainly financing energy projects. China's spending rose 20% to $67.7 billion, mostly owing to greater investment in solar energy. Policy changes such as expired subsidy programmes led to lower investment in the United States (down 32% to $44.2 billion), Italy (a drop of 51% to $14.7 billion) and Spain (a 68% reduction to $3 billion).

Arctic drilling

US secretary of the interior Ken Salazar announced a federal review of the offshore drilling programme in the Arctic on 8 January, focusing on a series of problems encountered by oil-and-gas giant Shell. Late last month, Shell lost control of a drilling rig, which ran aground in the Gulf of Alaska; the company briefly lost control of another rig in July. The review will focus on these issues as well as on challenges that the company has faced in implementing its oil-spill-preparedness plan. A report is scheduled for completion in early March.

↻ **NATURE.COM**
For daily news updates see:
www.nature.com/news

TREND WATCH

Human health and life expectancy are poorer in the United States than in other high-income nations, according to a 9 January report by the National Research Council in Washington DC. Compared with 16 peer countries, including France, Japan and the United Kingdom, the United States has some of the highest rates of infant mortality and obesity. The death rate in the United States is higher than average for most ages, with the largest disparities at younger ages (see chart).

POOR HEALTH IN THE UNITED STATES

Compared with 16 other rich countries, the United States had the highest death rates in most young age categories in 2005–09.

WOMEN — MEN
Mortality (%), log scale
Age (years)
— United States
— Other countries
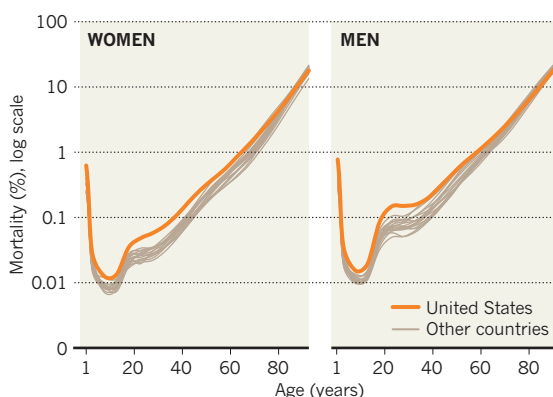
▶ Graham. "Can we get rid of that residual infection by adding an antibody, and would it work well enough to be cost-effective and logistically feasible?" he asks. The trial is designed to answer that question by giving the mother–baby pairs oral ARV therapy while half of the infants also receive monthly injections of VRC01. A control group would get a placebo.

The study will proceed only if VRC01 is approved as an investigational drug by the US Food and Drug Administration, and it would first be safety-tested in adults and infants in the United States. Investigators then hope to move to trial sites in Botswana, Malawi, South Africa, Tanzania, Uganda, Zambia and Zimbabwe. But the critics say that even if the antibody reduces the HIV infection rate, it is unlikely to be widely used in Africa. Whereas ARV therapy is delivered as oral tablets that cost less than US$200 per patient per year in poor and middle-income countries, the antibody must be injected by trained staff. In the United States, monoclonal antibody treatments commonly cost thousands of dollars per year.

Physician and AIDS expert David Ho of the Aaron Diamond AIDS Research Center in New York City is also planning to test an anti-HIV monoclonal antibody, but says that his group decided against running their trial in newborns after working in a poor area in southern China. There, he found that using ARV therapy as well as providing clean water and formula so that women do not have to breastfeed reduced mother-to-child HIV transmission to 1% (Z. Zhou *et al. J. Acquir. Immune Defic. Syndr.* **53** (suppl. 1)**,** S15–S22; 2010). With a transmission rate that low, Ho says, "you are not going to be able to do a meaningful study". Instead, he is planning to conduct his group's prevention study in men who have sex with men.

But Graham says that discouraging breastfeeding, with its many health benefits, is not an acceptable route for reducing infection risk. And Catherine Hankins, deputy director of science at the Amsterdam Institute for Global Health and Development, says that it is premature to worry about cost. She points out that the basic ARV tablets used in Africa dropped in price from more than $10,000 to as little $140 annually in the past decade. "To say up front that something is too expensive, forget it — there are a lot of things we wouldn't have today if people had said that." ∎

FINANCE

Insider trading sparks concerns

Universities indulge researchers' ties to finance industry.

BY HEIDI LEDFORD

It is as predictable as a heartbeat: when crucial medical results are at hand, cardiologist Jean-Luc Vachiery knows that the hedge-fund managers will come calling. Vachiery, who tests experimental therapies for a rare condition called pulmonary arterial hypertension at the Free University of Brussels, is privy to confidential information about clinical trials that could be valuable if used — illegally — to guide investments. As such, his conversations with the financial sector can be tense. "They tell you, 'don't say anything that's confidential,'" says Vachiery. "And then they ask you for confidential information."

Vachiery, like about 35,000 academics, is a member of the Gerson Lehrman Group, based in New York, one of about 40 'expert network' companies proliferating in the United States and elsewhere that connect clients, often from the financial industry, with experts who can provide technical information. The company recruits heavily from academia and was reluctantly thrust into the limelight last year when one of its most prestigious experts — Sidney Gilman, a neurologist then at the University of Michigan at Ann Arbor — admitted to tipping off a hedge-fund manager about clinical-trial data before they were made public. The result: US$276 million in illicit gains for the hedge fund, the largest insider-trading case the US Securities and Exchange Commission (SEC) has ever handled.

The case is just one in a string of SEC probes launched since 2009 into whether expert networks are trafficking insider information to the financial industry (see 'Trading on expertise'). That investigation has already led to charges against nearly 30 people with connections to expert networks, and is continuing to yield fresh targets, particularly in the health-care industry. It also highlights the tightrope that researchers walk when they consult for the financial industry, for rates that can reach $1,000 per hour. "The only reason anybody wants to talk to you from a financial company is for insider information," says Arthur Caplan, a bioethicist at New York University's Langone Medical Center. "That's the start of the story and that's the end of the story."
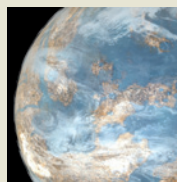
Despite the risks, few research institutions have enacted policies to limit staff consultations for the financial industry. Many universities, as well as the US National Institutes of Health, require investigators to disclose consulting income above a given threshold, but few distinguish financial consulting from consulting for drug companies. That's an important distinction, argues Eric Campbell, who studies conflicts of interest at Massachusetts General Hospital in Boston. "This is just a short-term bet for a very select group of people to make a whole lot of money," says Campbell. "This is not something that doctors or researchers should be involved in."

Even so, the networks and the experts they employ maintain that they are doing nothing illegal. Gerson Lehrman, like many other expert networks, has policies to prevent the exchange of confidential information, including a mandatory online training course, and pre-interview questionnaires to prevent experts with insider information from being matched with clients seeking information about that particular field. "All of this makes us a bad place to break the law,"

> *"This is not something that doctors or researchers should be involved in."*

---

says Alexander Saint-Amand, chief executive of Gerson Lehrman. Experts contacted by *Nature* say that they are often asked for mundane medical information, and are only sometimes asked to divulge secrets. All of them, including Vachiery, say that they never have. "Probably 99% of the use of expert networks has to be considered legitimate," says Michael Mayhew, chief executive of Integrity Research, an investment research firm in New York that tracks the expert network industry.

Yet Vachiery says that he is acutely aware of the techniques that sophisticated hedge-fund managers use to dissect his answers. Even as he deflects the questions in face-to-face meetings and teleconferences, he knows that his clients might gather small clues from his body language or his demeanour, patching those clues together with crumbs they've garnered from other experts. "With these bits and pieces they build their story," he says. "They even gauge your nonreaction."

With the risks so high, Mayhew says that he is surprised that hospitals and universities haven't banned participation in expert networks. "If you are afraid that sensitive information might get out, you probably should be proactive."

One of the few medical institutions to monitor financial consulting closely is the Cleveland Clinic in Ohio, which since 2005 has required that all such relationships be reviewed by a lawyer and is now considering an outright ban. Caplan says that professional associations should take a lead role in discouraging members from financial consulting. But Heather Pierce, a director of science policy at the Association of American Medical Colleges in Washington DC, says that she does not know of any professional societies that do.

Marcia Boumil, who oversees conflicts of interest at Tufts University Medical School in Boston, says that the school has no plans to create separate financial consulting disclosure requirements, nor does it they plan to limit researchers' ability to consult. "We're talking about highly respected professionals and we're not really in the business of policing what they do." Nevertheless, Boumil notes that her office might well have noted something odd about Gilman, had he worked at Tufts. Gilman earned $1,000 an hour for his consultations, and made more than $100,000 a year from them — two details that he would have been required to report at Tufts — and that would have raised a red flag.

But David Wazer, a radiation oncologist at Tufts, defends the $1,000 per hour consultations Gerson Lehrman arranges for him. "I'm worth it," he says, adding that, although he often does not know who his clients work for, he has never been asked to give up confidential information. The most he has ever made in one year is $4,000, "not the sort of compensation that would drive me to do anything peculiar", he says.

TRADING ON EXPERTISE

Since 2009, the US Securities and Exchange Commission (SEC) has been cracking down on insider trading by participants in expert networks.

20 NOVEMBER 2012 The SEC charges hedge-fund manager Mathew Martoma (pictured) with insider trading on confidential data gleaned from neurologist Sidney Gilman, whom he met through Gerson Lehrman.
Illicit gains: **US$276 million**.

17 FEBRUARY 2012 The SEC charges Oregon-based expert network firm Broadband Research Corporation and its owner John Kinnucan with trading inside information about 12 technology firms.
Illicit gains: **$110 million**.

3 FEBRUARY 2011 The SEC charges six consultants and employees of Primary Global Research in Mountain View, California, with illegally tipping off hedge funds about technology companies.
Illicit gains: **$6 million**.

2 NOVEMBER 2010 The SEC charges French hepatologist Yves Benhamou with revealing clinical-trial details to a hedge-fund manager he met through an expert network.
Illicit gains: **$30 million**.

16 OCTOBER 2009 The SEC charges Raj Rajaratnam and his hedge-fund advisory firm Galleon Management in New York of insider trading in the first of a series of cases that eventually implicated people with connections to expert networks.
Illicit gains: **$53 million**.

Instead, Wazer says that his motivation is to advance the development of radiation therapies by influencing investment. "The technologies that move forward are very much a function of where capital flows," he says. "And let's face it, these guys in financial firms determine the flow of capital." Others offer a different explanation: "The pay is good," says Donald Tsai, an oncologist at the University of Pennsylvania's Abramson Cancer Center in Philadelphia, who charges around $300 per hour and says that his clients typically ask him for "kindergarten-grade medical information". "It's fun and easy to do in my free time."

But in the wake of the SEC investigations, some researchers have decided that the stigma attached to expert networks outweighs

the pay. Mark Ratain, an oncologist at the University of Chicago in Illinois, severed ties with Gerson Lehrman when the SEC announced its investigation of expert networks. "I didn't want my name associated with this," he says.

And some of those who continue to participate say that they have become wary. Vachiery refuses repeat invitations from those who have prompted him for confidential information. Although he has enjoyed his conversations with some clients, the pressure has made him sour about the whole process. "I'm weighing each word I'm giving to these people and my level of trust is getting lower," he says. "Each time I receive a request now I'm thinking, 'Why am I doing this?'" ■ **SEE EDITORIAL P.271**
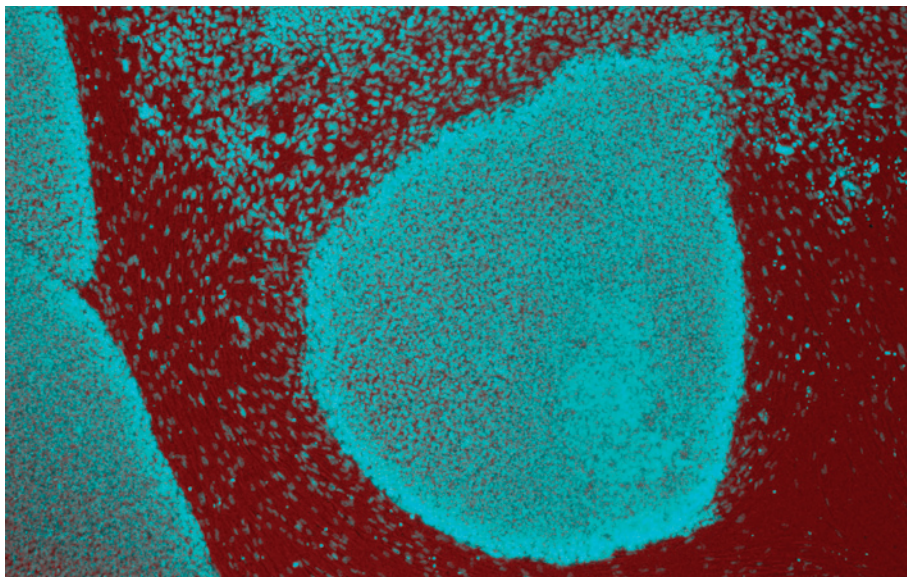
Researchers are keen to compare induced pluripotent stem cells (pictured) with their embryonic cousins.

STEM CELLS

Court lifts cloud over embryonic stem cells

But research on induced stem cells may be the real winner.

BY MONYA BAKER

The US Supreme Court's decision last week to throw out a lawsuit that would have blocked federal funding of all research on human embryonic stem cells cleared the gloom that has hung over the field for more than three years. Yet the biggest boost from the decision might go not to work on embryonic stem (ES) cells, but to studies of their upstart cousins, induced pluripotent stem (iPS) cells, which are created by 'reprogramming' adult cells into a stem-cell-like state.

At first glance, iPS-cell research needs no help. Researchers flocked to the field soon after a recipe for deriving the cells from adult mouse cells was announced in 2006, partly because this offered a way to skirt the thorny ethical issues raised by extracting cells from human embryos. But the real allure of iPS cells was the promise of genetically matched tissues. Adult cells taken from a patient could be used to create stem cells that would, in turn, generate perfectly matched specialized tissues — replacement neurons, say — for cell therapy. Although the number of published papers from iPS-cell research has not yet caught up with that of ES-cell work (see 'Inducing a juggernaut'), US funding for each approach is now roughly matched at about US$120 million a year.

But, as iPS cells crop up in ever more labs, ES cells — generally cheaper, better behaved and backed by an extra decade's worth of data — promise to have an important supporting role. Ever since iPS cells were described, researchers have been trying to understand just how similar they are to ES cells. iPS cells begin with different patterns of gene expression, and they can also acquire mutations during the reprogramming process, which means that every iPS cell must be thoroughly evaluated before it can be used in any study. "Human ES cells will always be the standard to which other cells will be compared," says Roger Pedersen, who studies how stem cells retain embryo-like states at the University of Cambridge, UK.

Federally supported ES-cell research was shut down in the United States on 23 August 2010, a year after a lawsuit was filed by two opponents of human ES-cell research, and remained frozen for more than two weeks. Many investigators shied away from the field for fear of having to shut down again. The Supreme Court's move has reassured investigators such as Candace Kerr, who studies early development of the brain at the University of Maryland School of Medicine in Baltimore. As a young scientist working towards tenure, she felt particularly vulnerable to the threat of ES-cell funding being stopped. So she switched to iPS cells in 2010, while the lawsuit was working its way through the US court system. With the litigation over, she says she need not "hesitate or fear" adding to her work with experiments using ES cells, which she finds much easier to prompt into neurons than iPS cells. "I am excited and relieved by this decision," she says.

ES cells do, however, have a head start in the clinic. Former heads of the biotech company Geron, based in Menlo Park, California, last week announced an agreement to acquire stem-cell assets including the company's flagship human ES-cell trial, in which precursors of neural support cells grown from human ES cells were injected into people with spinal-cord injuries. In 2011, Geron had shut down the $25-million-a-year effort, after the trial had gone on long enough to show that an ES-cell-based product could be safe. However, it had not shown signs of benefiting patients.

If revived, the spinal-cord trial would double the number of companies sponsoring human clinical trials for ES-cell therapies. At present, the only active trials are those sponsored by Advanced Cell Technology of Santa Monica, California: these are showing early evidence that a product made from human ES cells can help to rebuild the layer of cells that support photoreceptors in the eyes of people with certain types of blindness.

But iPS cells are edging towards the clinic too. Advanced Cell Technology says that it will begin talking to the US Food and Drug Administration this March about the safety studies required to test platelets derived from iPS cells in humans, and Japan is setting up a stem-cell bank of some 75 iPS-cell lines intended for future therapies. Even James Thomson at the University of Wisconsin–Madison, who isolated the first human ES cells in 1998, says that the future belongs to induced stem cells. His lab hasn't bothered to derive a single human ES-cell line since iPS cells were reported. "But we've derived dozens if not hundreds of human induced pluripotent stem cells," he says. ∎

INDUCING A JUGGERNAUT

The proportion of stem-cell publications and posters featuring induced pluripotent stem (iPS) cells rose rapidly in the years after their 2006 discovery.

- 🟥 iPS cells only
- 🟩 Both embryonic stem (ES) and iPS cells
- 🟦 ES cells only
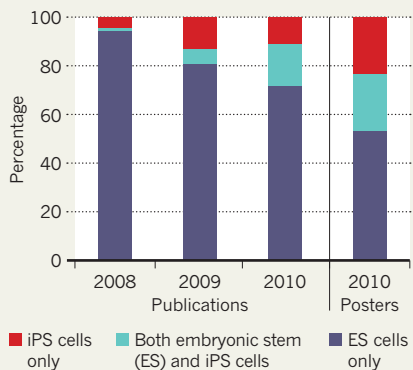
TECHNOLOGY

Quantum dots go on display

Adoption by TV makers could expand the market for light-emitting nanocrystals.

BY KATHERINE BOURZAC

Live from your living room, in super-saturated colour: it's the quantum-dot TV! Researchers working with nanoscale fluorescent particles called quantum dots have long predicted groundbreaking achievements, such as ultra-efficient light-emitting diodes (LEDs) and solar cells, but the technology has found mainly niche applications. That could change with the announcement last week that QD Vision, based in Lexington, Massachusetts, would supply Sony Corporation of Tokyo with quantum dots for flat-screen televisions that will transmit more richly coloured images than other TVs on the market.

Demand for quantum-dot displays, say industry watchers, could benefit quantum-dot companies, bring down the price of these nano-materials and boost other applications that have stalled. "Displays are a potential market that could help quantum-dot companies find traction," says Jonathan Melnick, an analyst at Lux Research in Boston, Massachusetts.

Quantum dots are crystals about 10 nanometres in diameter, made from a semiconductor material, commonly cadmium selenide. They are so tiny that their shape and size affect the quantum properties of their electrons, in particular their energy gap — the energy needed to kick electrons into a higher-energy band — which determines the colour of light that the material can emit. Whereas a bulk semiconductor is limited to emitting a single colour of light, researchers can tune the precise colour a quantum dot will absorb and re-emit by tailoring its size.

Discovered in 1981, quantum dots did not find applications until 2002. That was when the Quantum Dot Corporation of Hayward, California, began selling them to cell biologists, who prize them as fluorescent imaging labels for proteins and other biological molecules. As recently as 2010, the biomedical sector was responsible for US$48 million of $67 million in total quantum-dot revenues, according to BCC Research of Wellesley, Massachusetts.

Quantum dots have shown promise for electronics, too — for example in solar cells, in which a mix of quantum dots tuned to absorb different wavelengths of light could capture more of the energy in the solar spectrum. But one hurdle to their exploitation was their temperature sensitivity. Near the backlight of a liquid-crystal display (LCD), for example, temperatures can be around 100 °C. At this temperature, the dots lose efficiency and up to half

Sony has announced that it will embed quantum dots in its latest flat-screen televisions.

of their brightness, says QD Vision co-founder and chief technology officer Seth Coe-Sullivan. He says that the company spent a long time tuning the chemistry of its quantum dots to make them stable at higher temperatures.

Moungi Bawendi, a chemist at the Massachusetts Institute of Technology in Cambridge and a co-founder of QD Vision, admits that the company also made some business miscalculations. For its first product, in 2009, it provided Nexxus Lighting of Charlotte, North Carolina, with quantum-dot coatings to convert the harsh glare of LEDs into a warmer glow, to make them more appealing as long-life, low-energy light bulbs. But Bawendi says that LED designs and technology for the light-bulb market evolved too fast for the quantum-dot coatings to keep up. "You design something, and six months later it doesn't work," Bawendi says. "Television technology is more stable."

His optimism will be tested this spring with the company's quantum-dot debut in Sony LCD televisions, to be sold under the Triluminos brand name. The contrast with today's flat screens begins with the light source. Conventional LCDs use a high-intensity blue LED backlight whose glow is converted by a phosphor coating to create a broadband, white light used to make the moving TV images. The new Triluminos televisions instead pair an uncoated blue LED with a thin glass tube filled with quantum dots. Two kinds of quantum dots in the tube absorb some of the blue light from the backlight and re-emit it as pure red and green light. The resulting white light is

more intense at the wavelengths of these three specific colours than the white light made by a phosphor-coated LED, so that more colour comes through in the images.

Another quantum-dot company, Nanosys of Palo Alto, California, is providing 3M of St Paul, Minnesota, with material for a similar product. 3M will make a polymer film seeded with quantum dots that does the same job as QD Vision's glass tube. The film is layered between the LCD's stack of light filters, diffusers and polarizers, and similarly converts raw blue light into white light made up of pure colours. Nanosys and 3M announced their partnership in June 2012, but have not yet named any customers.

BCC predicts that, by 2015, optoelectronics, including display components, will make up $310 million of a total $666 million in quantum-dot revenues. Melnick says that these numbers might be overly optimistic, because quantum dots remain expensive. "Even on the low end, they still cost in the hundreds of dollars per gram, and range up to $10,000 per gram," he says. But demand from 3M and Sony could help to bring prices down. Although neither QD Vision nor Nanosys would comment on the volume of material they expect to make this year, or their selling price, both say that they are scaling up their manufacturing volume.

Bawendi is not surprised that it took quantum dots so long to find their footing. "You could argue that 30 years is about the right amount of time from fundamental discovery to applications," he says. ∎

STEVE MARCUS/REUTERS

GENETICS

Behaviour genes unearthed

Speedy sequencing underpins genetic analysis of burrowing in wild oldfield mice.

BY EWEN CALLAWAY

Even as home experiments go, Hopi Hoekstra's one was peculiar: she built a giant plywood box in her garage in San Diego, California, filled it with more than a tonne of soil and then let a pet mouse dig away.

"This thing was bursting at its seams and held together with duct tape," says the evolutionary biologist, now at Harvard University in Cambridge, Massachusetts. "But it worked." It allowed her to study the genetics of burrowing behaviour in a controlled setting. Armed with plastic casts of the burrows and state-of-the-art sequencing, Hoekstra's team discovered clusters of genes that partly explain why the oldfield mouse (*Peromyscus polionotus*) builds elaborate two-tunnel burrows, whereas its close relative, the deer mouse (*Peromyscus maniculatus*), goes for a simple hole in the ground[1].

The findings highlight an underappreciated benefit of a genomics revolution that is moving at breakneck speed. Thanks to cheap and quick DNA sequencing, scientists interested in the genetics of behaviour need not limit themselves to a handful of favourite lab organisms. Instead, they can probe the genetic underpinnings of behaviours observed in the wild, and glean insights into how they evolved. "In my mind, the link between genes and behaviour in natural populations and organisms is the next great frontier in biology," says Hoekstra.
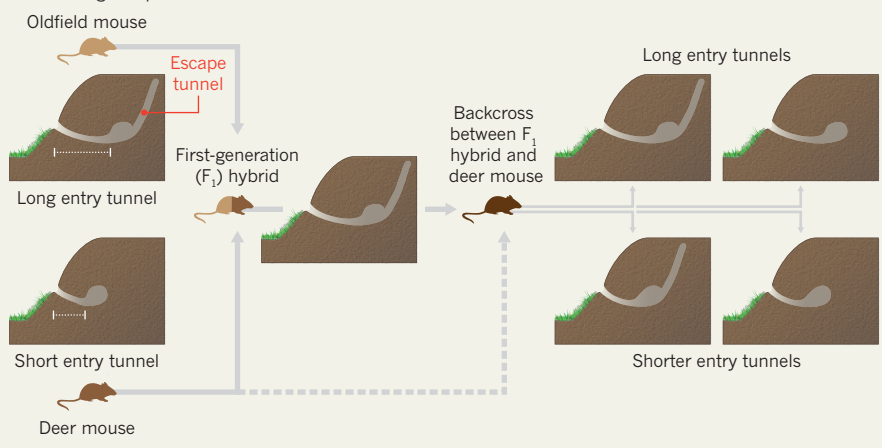
Oldfield mice are native to the southeastern United States, where they burrow in soils ranging from sandy beaches to silt-rich clays. Wherever they dig, their holes look much the same, with a long entrance tunnel and a second tunnel that stops short of the surface and allows them to escape predators. Such invariability hints that the trait is encoded in DNA, says Hoekstra.

To find out where, she and her Harvard colleagues Jesse Weber and Brant Peterson cross-bred oldfield mice with deer mice, whose burrows are shallow and lack escape routes. The offspring continued to build complex tunnels, suggesting that the oldfield burrowing genes were dominant (see 'The genetics of burrowing').

A second round of breeding between the first-generation crosses and deer mice revealed that genes linked to burrow length were distinct from those influencing the escape tunnel. Some offspring produced short tunnels with escape routes, whereas others produced long tunnels without them. DNA analysis revealed that three genetic regions are responsible for much of the variation in tunnel length, and a fourth affects escape-tunnel digging.

THE GENETICS OF BURROWING

Mouse cross-breeding shows that distinct gene clusters control the length of entry tunnels and whether or not the mice dig escape tunnels.

Oldfield mouse — Escape tunnel — Long entry tunnel — Short entry tunnel — Deer mouse — First-generation (F₁) hybrid — Backcross between F₁ hybrid and deer mouse — Long entry tunnels — Shorter entry tunnels

PAUL JACKMAN/NATURE

"This paper is awesome," says Cornelia Bargmann, a neurogeneticist at Rockefeller University in New York, noting that it combines cutting-edge molecular-genetics tools with established cross-breeding techniques to study behaviours that have been observed for more than a century in the wild. In the past, geneticists interested in unravelling behaviour had to focus on lab animals for which mutant and transgenic strains and genetic data were available, she says. "But there were always questions we knew would be more interesting in wild animals." Bargmann and her team studied various wild strains of *Caenorhabditis elegans* flatworms to identify genes and brain circuits involved in seeking out new sources of food[2].

"The link between genes and behaviour in natural populations is the next great frontier."

Cheap genome sequencing is the foundation of this new behavioural genetics. In a paper published on *Nature*'s website this week[3], a team led by Laurent Keller at the University of Lausanne, Switzerland, sequenced hundreds of fire ants (*Solenopsis invicta*) and identified a region that explains why some colonies have only one queen whereas others tolerate more. The region contains hundreds of genes, and Keller thinks that vast behavioural differences among other insects might be explained by such 'social chromosomes'.

These studies offer clues to the evolution of behaviours seen in the wild, says Catherine Peichel, an evolutionary geneticist at the Fred Hutchinson Cancer Research Center in Seattle, Washington. In unpublished work, Peichel and her postdoc Anna Greenwood have discovered a gene that explains differences in schooling behaviour in stickleback fish. The marine variety swims in tight schools, whereas its freshwater cousin swims solo. Such behaviours were probably selected for by different environments — predator-filled seas with few hiding places versus vegetation-filled streams. And Hoekstra says her findings suggest that complex burrowing in mice evolved piecemeal, by combining the genetic modules responsible for simpler digging behaviours.

Hoekstra's team has already found that the key regions for burrowing include genes involved in addictive behaviour. "When you watch these mice, you can't help but think they look like they're addicted to burrowing," Hoekstra says. "You drop one of these oldfield mice in one of our burrowing boxes and they immediately start digging."

The ultimate proof of any gene's role will be to engineer deer mice to build complex burrows by inserting oldfield mice genes, says Hoekstra. Those experiments would have been unthinkable a few years ago, when genome engineering was possible in only a handful of mammals, but this is now conceivable, says Bargmann (see *Nature* **455,** 160–164; 2008). ∎

1. Weber, J. N., Peterson, B. K. & Hoekstra, H. E. *Nature* **493,** 402–405 (2013).
2. Bendesky, A., Tsunozaki, M., Rockman, M. V., Kruglyak, L. & Bargmann, C. I. *Nature* **472,** 313–318 (2011).
3. Wang, J. *et al. Nature* http://dx.doi.org/10.1038/nature11832 (2013).

More children will be receiving injected killed polio vaccine in a bid to finally eradicate the virus.

HEALTH

Vaccine switch urged for polio endgame

Inactivated virus vaccine could deliver the final blow.

EWEN CALLAWAY IN KADUNA CITY, NIGERIA

By sunrise on a warm December morning, Janila Shulu's team are out in the dirt roads and alleyways of Ungwan Rimi, a poor neighbourhood in a predominantly Muslim section of Kaduna city in northern Nigeria. Three female health workers, accompanied by a community leader, dart from house to house, squeezing a few drops of polio vaccine into the mouths of all the young children they can find, even those who pass by on the street. By 1 p.m., after giving hundreds of doses, they stop for the day — the first of a national five-day effort.

Such campaigns are the backbone of the global push to eradicate polio, but this month the World Health Organization (WHO) in Geneva, Switzerland, proposed a shift in vaccination strategy from oral vaccines to injected ones that may have to be administered in clinics. The change is needed to mop up the last remaining pockets of polio, but experts say that it poses challenges in places such as Kaduna city, which have poor access to health care.

The new policy is an important step towards eradication, says Nicholas Grassly, an epidemiologist at Imperial College London, but implementing it will be difficult. "There are some big ifs as to whether it can happen," he says.

Jonas Salk is credited with developing the first polio vaccine in 1955, an injected vaccine containing killed virus, but the oral live vaccine devised a little later by his competitor Albert Sabin is the workhorse of the Global Polio Eradication Initiative. This public–private effort, started in 1988 and coordinated by the WHO, has cost about US$8 billion so far. The Sabin vaccine is composed of three live but crippled strains of polio. It is cheap and easy to administer, making it ideal for national campaigns that involve tens of thousands of minimally trained workers.

But the live viruses in the Sabin vaccine can revert to disease-causing forms, especially in populations where immunity is not widespread. Northern Nigeria has been battling such vaccine-derived outbreaks since 2005, and one emerged last year in Pakistan (see *Nature* **485,** 563; 2012).

In a 4 January announcement, the WHO called for oral polio vaccine containing the polio strain type 2, one of the Sabin vaccine strains, to be phased out — perhaps in as little as two years. The wild form of type 2 has been stamped out globally, but vaccine-derived forms of the strain still circulate in Nigeria and neighbouring countries. Oral polio vaccination will continue, but it will use a vaccine that protects against just the two other types of polio virus that are still circulating in their wild form in Nigeria, Pakistan and Afghanistan.

Meanwhile, the policy also calls for the introduction, as quickly as possible, of the oral vaccine's old competitor: the inactivated Salk vaccine. That costs more than ten times as much as the oral vaccine and requires trained health workers to administer it, says Roland Sutter, a vaccinologist at the WHO. But it carries no risk of causing polio. By giving children an inactivated vaccine that protects against all three subtypes of polio, health workers hope to gradually stamp out vaccine-derived outbreaks.
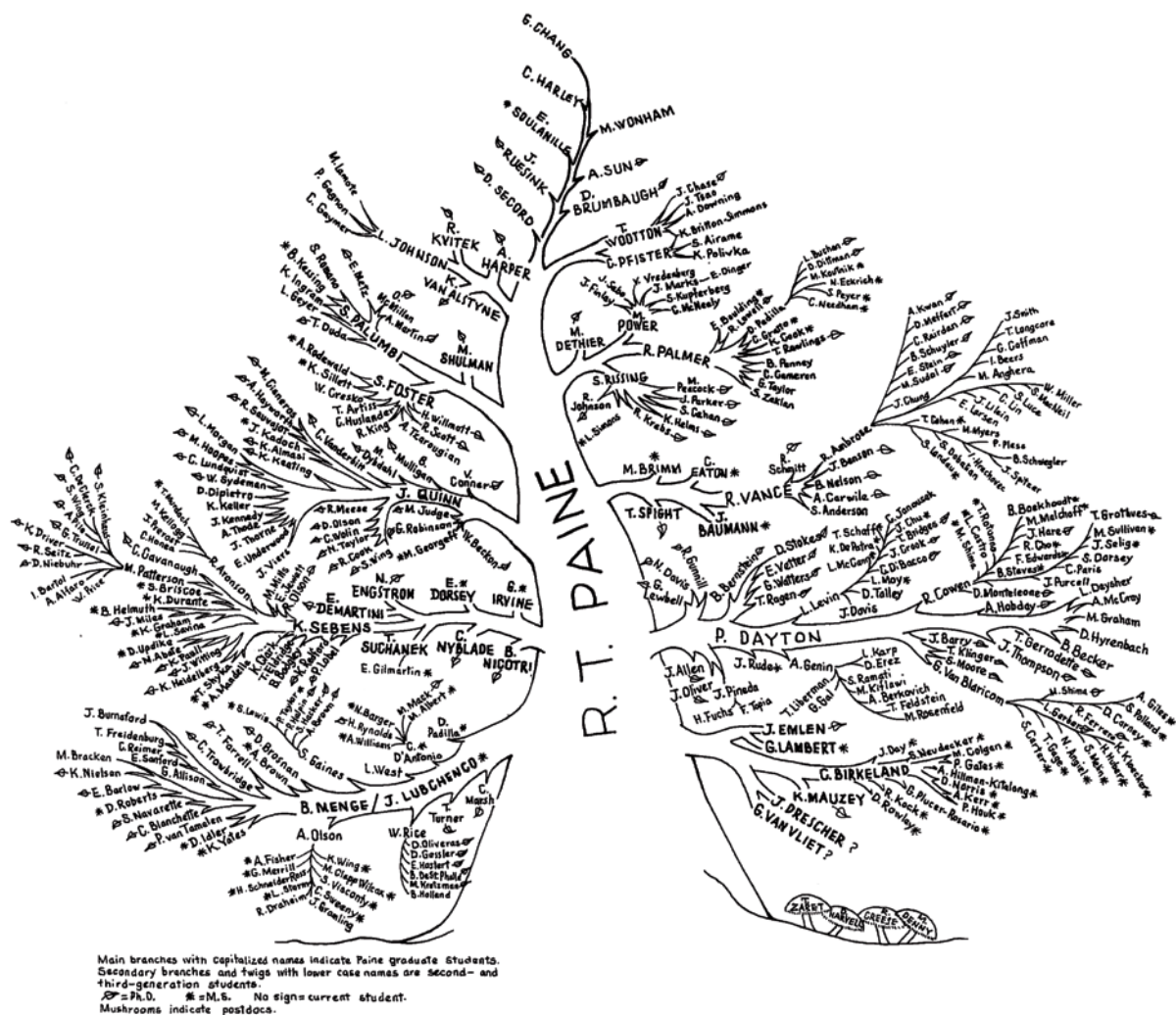
"You have to have a transition period" in which both oral and inactivated vaccines are used, "because if you stop cold turkey you're going to have outbreaks", says Vincent Racaniello, a virologist at Columbia University in New York city. Once the remaining wild polio types are wiped out, the WHO will phase out all oral polio vaccines.

The high cost of the inactivated polio vaccine remains a significant hurdle for the plan, which depends on a reduction in cost to less than 50 cents per dose from the current cost of more than $2, says Sutter. Boosting the immune response by including adjuvants, and delivering the vaccine under the skin instead of into muscle, could help to lower the dose required and cut costs, as could new kinds of vaccine, he says.

Health infrastructure poses another big hurdle, says Grassly. Delivering the vaccine in clinics instead of door to door will pose a challenge for Nigeria, which has one of the lowest rates of routine immunization in the world. Less than 50% of children receive a complete schedule of childhood vaccinations, and in parts of northern Nigeria that figure is around 10%.

"We as a global community have to do a much better job of integrating polio and routine immunization," says Zulfiqar Bhutta, an immunization expert at Aga Khan University in Karachi, Pakistan, and a member of the WHO committee that issued the new vaccination policy. He sees the eventual switch to inactivated vaccines as an opportunity to align polio eradication with routine immunization. "We should have done this a lot earlier," he says. ∎

Main branches with capitalized names indicate Paine graduate students.
Secondary branches and twigs with lower case names are second- and
third-generation students.
🖋 = Ph.D.   ✳ = M.S.   No sign = current student.
Mushrooms indicate postdocs.

DYNASTY

Bob Paine fathered an idea — and an academic family — that changed ecology.

BY ED YONG

Bob Paine is nearly 2 metres tall and has a powerful grip. The ochre sea star, however, has five sucker-lined arms and can span half a metre. So when Paine tried to prise the creatures off the rocks along the Pacific coast, he found that his brute strength simply wasn't enough. In the end, he resorted to a crowbar. Then, once he had levered the animals up, he hurled them out to sea as hard as he could. "You get pretty good at throwing starfish into deeper water," he says.

It was a ritual that began in 1963, on an 8-metre stretch of shore in Makah Bay, Washington. The bay's rocky intertidal zone normally hosts a thriving community of mussels, barnacles, limpets, anemones and algae. But it changed completely after Paine banished the starfish. The barnacles that the sea star (*Pisaster ochraceus*) usually ate advanced through the predator-free zone, and were later replaced by mussels. These invaders

Bob Paine's academic family tree, drawn by Marian Kohn in 1999.

crowded out the algae and limpets, which fled for less competitive pastures. Within a year, the total number of species had halved: a diverse tidal wonderland became a black monoculture of mussels[1].

By re-engineering the coastline in this way[1], Paine dealt a serious blow to the dominant view in ecology of the time: that ecosystems are stable dramas if they have a diverse cast of species. Instead, he showed that individual species such as *Pisaster* are prima donnas, whose absence can warp the entire production into something blander and unrecognizable. He described these crucial creatures, whose influence far exceeds their abundance, as keystone species, after the central stone that prevents an arch from crumbling. Their loss can initiate what Paine would later call trophic cascades — the rise and fall of connected species throughout the food web. The terms stuck, and 'keystone' would go on to be applied to species from sea otters to wolves, grey whales and spotted bass.

Today, ecology students take these concepts for granted — but they shook the field when Paine first articulated them in the 1960s. "He's been one of the most influential ecologists in the last half century," says Simon Levin, a mathematical ecologist at Princeton University in New Jersey, and one of Paine's closest friends. The revelation that not all species are equal was as disruptive to ecology as the loss of *Pisaster* was to Makah Bay. So was Paine's insistence on tinkering with nature — what some have called kick-it-and-see ecology — at a time when most ecologists simply observed it.

But Paine — an organism whose disproportionate influence equals that of any starfish or sea otter — has also changed the ecosystem of scientists. In his five-decade career, he has trained a thriving dynasty of around 40 students and postdocs, many of whom are now leading ecologists themselves and who consider their time with Paine formative. They include Paul Dayton at the Scripps Institution of Oceanography in La Jolla, California, who has shaped understanding of rocky shores, kelp forests and Antarctica's sea floor; Bruce Menge at Oregon State University in Corvallis, who expanded Paine's research to coasts worldwide; Jane Lubchenco, who heads the US National Oceanic and Atmospheric Administration (NOAA) in Washington DC; and Steve Palumbi at Stanford University in California, who used genetics to track the illegal trade of whale and dolphin meat.

"There are other ecologists as famous as Bob, but if you look at their list of students, there aren't nearly as many whom you know by reputation," says Chris Harley at the University of British Columbia in Vancouver, Canada, who was one of Paine's most recent PhD students. Once Paine's students' students are taken into account, his academic family easily stretches into the hundreds. "Everyone is tied to Paine," says Craig McClain, a deep-sea biologist at the National Evolutionary Synthesis Center in Durham, North Carolina, who is three links removed — his postdoctoral adviser was Dayton's student.

FAMILY VALUES

Science hosts many such dynasties: successions of academic leaders related not by blood, but by mentorship. Each generation inherits attitudes, philosophies and technical skills from the one before. Some, like Paine's, are particularly fertile, sprouting lush branches on the academic tree and driving a field in a new direction. But Paine's dynasty is remarkable not just for its scientific influence, but for its dedicated, tight-knit nature. Thanks to Paine's original — and widely applicable — ideas, his emphasis on independent thought by his protégés and his fun, irreverent nature, almost every member has stayed in science, and specifically in ecology or marine biology.

"It's a surprising list of superstars — great mentors of graduate students, who have published interesting work," says Paine, who retired in 1998 but is still active in the field. These days, Paine can be spotted at ecological meetings by the swarm of academic descendants milling around him. Perhaps in this rich family, there are lessons about why some scientific dynasties flourish and grow, whereas others never bud.

Paine's name is synonymous with coastal life, but his introduction to natural history began on terra firma. As a child in Massachusetts, he went on regular birdwatching walks with a neighbour, who insisted that he record everything he saw. "That was extraordinarily good training," says Paine; it instilled an appreciation for nature and careful observation.

After studies at Harvard University in Cambridge, Massachusetts, and two years' military service, he got his PhD in zoology from the University of Michigan in Ann Arbor, under the late Fred E. Smith. A qualified ecologist in search of an ecosystem, Paine started his own group at the University of Washington in Seattle in 1962 and embarked on his seminal work in Makah Bay. But despite its abundance of starfish, the bay was plagued with marauding beachcombers, who would tamper with Paine's experiments. He craved isolation.

He found it in 1967, after a salmon-fishing trip in the Pacific Ocean, when on a whim he landed on a small island called Tatoosh, just off Washington's Olympic Peninsula. "I just sat around and gawked," he says. "The place was a wonderland of pattern." He saw hundreds of species jostling for space and dominance. He saw starfish prising open mussels, and predatory sea snails drilling into barnacles. He saw whales and sea lions swimming offshore, while raptors and seabirds hovered overhead. And he saw … no one else. With the island protected by the local Makah tribe, there was no one around to ruin his experiments. "I said: this is where I'm going to work."

ISLAND RETREAT

Paine repeated his starfish-clearing experiment on Tatoosh, and saw the same events unfold on a bigger scale: the loss of *Pisaster* triggered a black landslide of mussels that crushed its way down 40 metres of coast. Elsewhere, in a zone dominated by brown algae, Paine systematically added or excluded seven grazing animals by building rings of paint and putty[2]. Two of the species — a sea urchin and a mollusc called a chiton — would annihilate the algae when present in high numbers. The others, all molluscs, had no effect. "It was a colossal effort," says Paine, and the first time that anyone had quantified the influence of so many species in a community. It showed that most are weak interactors, whose absence goes unnoticed. Only a few — including keystone species — are strong interactors that can radically reshape their world. "It was a starting place for untangling the complexity of interactions," says Paine. "If all species were created equal, you wouldn't know where to start."

Paine's views on experimental ecology rapidly percolated through the field, and attracted waves of eager students. "It was a pioneering stage in the field," says Menge. "We felt that we were really the first ones to be doing these sorts of experiments." Tatoosh gave them a place to experiment, and most of Paine's protégés have done tours on its weather-beaten terrain. They camped in garage-sized buildings, abandoned from the island's days as a coastguard outpost. There was no running water and, until a decade ago, no electricity. "It was a brutal environment, and I was out there suffering with them," says Paine. He kept the atmosphere spartan, splashing out only on heavy-duty clothing and boots — the "Paine lab uniform".

Unlike many principal investigators, who dole out predefined projects to new recruits, Paine encouraged freedom and individuality. "He was hands-off until it was necessary to be hands-on," says Menge. Students walked to disparate corners of the island to explore their own passions. One tagged larval rockfish; another studied barnacle dynamics. In the evenings, the group traded data over campfires. "You'd have a day of working the shoreline, you'd trudge back up to the island, and he invariably asked: What did you learn today?" says Tim Wootton from the University of Chicago, Illinois, who studied under Paine in the 1980s. Paine set a high intellectual bar, and was gruff and challenging. "He's very brusque and abrupt, and he kind of scowls a lot," says Lubchenco. But he also encouraged irreverence and mischief — he once signed the name of one his students, Peter Kareiva, on Christmas cards full

"All my students were smarter than me but just less knowledgeable."

of lascivious limericks and posted them to the world's top ecologists.

Paine treated his students as peers, supporting their endeavours without directing them. (Smith had treated Paine the same way.) When they published, he kept his name off their papers unless he had had a heavy hands-on role in the research — an ethos that seems unthinkable today. "It hurt him a bit," says Kareiva, now one of Paine's closest friends and chief scientist for the Nature Conservancy, an environmental organization based in Arlington, Virginia. "His presence in the literature would be ten-fold if he hadn't done that." Harley adds: "If Bob Paine were reborn into the system as a junior professor right now, he would turn out excellent graduate students but I don't know how many grants he'd get."

Soon, Paine's students were growing up and embarking on careers of their own. Few have spawned as rich a legacy as Jane Lubchenco and Bruce Menge. They met as graduate students in Paine's lab in 1969, married two years later and began a partnership that has generated

Lubchenco and Menge wanted to probe other environmental impacts on coastal ecosystems, and so began a decades-long world tour that took Paine's experimental approaches to Chile, South Africa and New Zealand. Their studies showed[4], for example, that the intermittent upwelling of nutrient-rich water can intensify competition, predation and other interactions between marine species.

But they also moved — to Paine's disapproval — from his solitary style to large teams. Their grandest venture is the Partnership for Interdisciplinary Studies of Coastal Oceans (PISCO) — a 13-lab collaboration that began in 1999. The participants run regimented studies along a 1,900-kilometre stretch of coastline from Alaska to Mexico, including a census of local marine life and measurements of water conditions. There are some Paine-style manipulations, but PISCO is more about collecting data to reveal how the oceans are changing. When Oregon's beaches started filling with dead crabs and fish in 2002, PISCO's data quickly

Bob Paine with starfish on Tatoosh island, Washington, in 2011 (left); Jane Lubchenco and Bruce Menge with students Eric Sanford and Tess Freidenburg, 1997.

more than 31 students and 19 postdocs. After the pair left Paine's lab, they took his experimental approach to the US east coast; she focused on plants and herbivores, while he concentrated on predators. By enclosing, excluding and removing species at different points along the New England shore, they showed[3] that fierce waves can keep predators such as starfish at bay, allowing mussels to dominate. But in sheltered areas, predators kept mussels under control, allowing Irish moss (*Chondrus crispus*), a type of red alga, to take over. The work revealed how the environment can control interactions between species.

GROWING ECOSYSTEM

After divvying up the New England food web between them, Lubchenco and Menge also split a tenure-track faculty position when they returned west to Oregon State University in 1976. This move, unprecedented at the time, allowed them to further their careers while spending time with their growing family. (One of their children, Duncan Menge, accompanied the Paine clan on field trips and this year is setting up his own ecology lab at Columbia University in New York — a blood child of the academic family.)

The students of the cheerfully described "Lubmengo lab" benefited from the yin-yang qualities of their mentors. Menge was impulsive and had an open-door policy. Lubchenco was deliberate and required appointments. Today, he is laid-back and wears Hawaiian shirts; she is intense and sharp-suited. "The combination of those two was really powerful," says Steve Gaines from the University of California, Santa Barbara, who was one of their star students. "You got an educational experience that would be hard for any individual to give you."

revealed the cause — an oxygen-deprived dead zone that had been mysteriously growing in the water off the coast (see *Nature* **466,** 812–814; 2010). "The problems we're trying to solve in ecology are way beyond the local scale," says Menge. "My dream is for PISCO-like entities in all the marine ecosystems around the world."

But that is not Paine's dream: big science is an anathema to him. "He was always grousing about how this massively expensive work was really not much more than stamp-collecting," says former student Richard Palmer, now at the University of Alberta in Edmonton. Paine fears that PISCO's approach will yield broad, unclear trends rather than detailed insights. Worse, it risks robbing students of the playful creativity that he tried to instil. "They're so involved with making the same measurements up and down the same kilometres of coast," he laments. "My loosey-goosey attitude to Tatoosh was to get brilliant people to do what they want to do." But Lubchenco and Menge argue that ecology's challenges, from global warming to ocean acidification, are outpacing the rhythms of solo experiments. "Bob fought that a lot," says Lubchenco. "People had to rebel against their parent."

Lubchenco's move into policy was similarly rebellious. She served as president of the American Association for the Advancement of Science in 1997 and of the Ecological Society of America between 1992 and

1993. And as administrator of NOAA from 2009, she has influenced the US government's response to the Deepwater Horizon oil spill, and its strategy to manage the oceans sustainably. "The culture we grew up with was very anti being relevant," says Lubchenco, who will leave NOAA at the end of

February to return to research. "I knew Bob didn't approve and I did it anyhow. It was really painful to rebel."

Any such disapproval has long since vanished, and Paine now speaks of Lubchenco's influence with deep respect. He has reluctantly become a key player in conservation himself, leading panels of scientists who have assessed the decline of the Steller's sea lion (*Eumetopias jubatus*) and the recovery of Alaskan waters following the 1989 *Exxon Valdez* oil spill. "My role was simply to oversee the mob," he says. "I was bullied into it."

NEXT GENERATION

Lubchenco's academic children — Paine's grandchildren — have been steeped in 'relevance' from the start. "We were all surrounded by this environment where it was encouraged to think about steps connecting the science to policy," says Heather Leslie at Brown University in Providence, Rhode Island, a former Lubchenco student.

work in drug metabolism and Julius Axelrod, who was awarded a Nobel prize for his work on the release of neurotransmitters — helped to revolutionize human pharmacology and neuroscience. "Their approach was: Don't feel you have to have all your i's dotted and t's crossed. Just go and do it," says Robert Kanigel, a writer who chronicled the dynasty in his book *Apprentice to Genius* (Macmillan, 1986).

But the Shannon lineage was radically different from Paine's. There was no premium on independence; instead, students served as apprentices to their masters. "In many of the key experiments, the younger person was doing the work of the mentor, and that sometimes led to resentments," says Kanigel. Such relationships can breed envy and mistrust, he says. "People wonder if they'll get the credit they deserve."

Paine proves that the opposite strategy works. "Treat your graduate students as human beings and be accessible to them. That sense of social equality is very important," he says. "All my students were smarter than

Paine (front row, centre) surrounded by colleagues and protégés at his retirement party in 1999 (left); Paine's former student Cathy Pfister on Tatoosh in 2012.

Gaines epitomizes this approach. The ultimate blend of the Paine and Lubmengo heritage, he is an experimental ecologist who uses basic science to influence policy, is not shy of collaborations and has won a national award for mentoring. "Steve sits right at that sweet spot in many different ways," says Lubchenco. Whereas Paine studies how species interact on individual shores, Gaines looks at connections across entire oceans. He has shown[5] that the larvae of fish and other marine life sometimes drift for hundreds of kilometres on ocean currents, even if the adults are fixed in place. "When Bob removed the sea star and got an explosion of mussels, those mussels were coming from somewhere else," he says. These planktonic drifters connect disparate parts of the ocean, with huge implications for marine reserves, where fishing is forbidden. Rather than reducing fish catches, Gaines showed[6] in 2005 that these reserves have the potential to make nearby fisheries more productive as their larvae disperse and replenish the stocks.

Gaines' own students, in the style of the entire family, have branched out into new areas. Kate Smith, who did her PhD with him and is now at Brown, applied Gaines' ideas on dispersing larvae to understand how infectious diseases disperse. Her ecological interests are the same as those of her academic parents and grandparents — where organisms are, why they are there and how they influence each other — but applied to microbes and continents, rather than invertebrates and tidal pools. "It's all related to connectivity," says Gaines.

There are other ways than Paine's to spawn a dynasty. Take, for example, the chain of biological luminaries that emerged in the 1940s, beginning with physiologist James Shannon at the US National Institutes of Health. His descendants — including Steve Brodie, who did pioneering

me but just less knowledgeable." This attitude selected for self-propelled, passionate students who could find their own way. "It helped in getting the experience to implement your own research programme," says Wootton.

Now aged 79, Paine is still conducting research and inspiring students. He joined a crop of them last summer for a week in Patagonia. "He'd be in the intertidal on his hands and knees and asking questions," says Kareiva, who was also on the trip. Paine also makes regular excursions to Tatoosh, where research is now run by Wootton and his wife, Cathy Pfister, who is also at the University of Chicago. Paine's mind is as sharp as ever, but having lost good binocular vision, his steps are less sure. "I hire my very athletic daughter to haul me around the rocks," he says. "It's not sufficient, but I can continue to keep track of my long-term experiments."

In 1995, Paine finally allowed starfish to return to the Tatoosh shore after 25 years of continuous exile. "I thought, I'm not going to live forever, so let's see what happens if I let the starfish back." Their numbers have rebounded, the mussels are losing ground, and the shore is returning to the state it was in before Paine's interference. The same cannot be said for ecology. It will never be the same after Paine. ■

Ed Yong *is a freelance writer and author of the blog Not Exactly Rocket Science, based in London.*

1. Paine, R. T. *Am. Nat.* **100,** 65–75 (1966).
2. Paine, R. T. *Nature* **355,** 73–75 (1992).
3. Lubchenco, J. & Menge, B. A. *Ecol. Monogr.* **48,** 67–94 (1978).
4. Menge, B. A. *et al. Ecol. Res.* **17,** 1–16 (2002).
5. Kinlan, B. P. & Gaines, S. D. *Ecology* **84,** 2007–2020 (2003).
6. Gaylord, B., Gaines, S. D., Siegel, D. A. & Carr, M. H. *Ecol. Appl.* **15,** 2180–2191 (2005).

FALLOUT OF FEAR

AFTER THE FUKUSHIMA NUCLEAR DISASTER, JAPAN KEPT PEOPLE SAFE FROM THE PHYSICAL EFFECTS OF RADIATION — BUT NOT FROM THE PSYCHOLOGICAL IMPACTS.

BY GEOFF BRUMFIEL

The first thing Kenichi Togawa does when he comes home from work is switch on his video-game console. The 39-year-old father of three spends hours each evening playing video games and drinking *shochu*, a strong Japanese liquor. He often falls asleep in front of the television, then wakes up shivering and crawls into bed with his wife, Yuka.

For nearly two years, Kenichi and his family have been refugees from the worst nuclear disaster in 25 years. On 11 March 2011, a giant earthquake struck off the northeast coast of Japan, sending a 13-metre-high wall of water into the Fukushima Daiichi nuclear power station and triggering meltdowns in three of the six reactors. The next day, just hours before the Unit 1 reactor exploded, the Togawa family fled their home 10 kilometres from the plant. Today, they live in a tiny flat outside the evacuation zone — one of dozens in a series of slate-grey temporary buildings in the northeast section of Fukushima prefecture. The five Togawas are bundled into three rooms totalling just 30 square metres, with windows poorly insulated against the winter winds.

The past 18 months have taken a mental toll on the family. Kenichi, who had worked at the nuclear plant, was once a keen judo fighter who went out often with friends, but the radioactivity has scattered his martial-arts club. These days, he exercises less and rarely socializes. He drinks more and has put on weight.

Yuka is prone to public outbursts of anger, unusual among Japanese women in the relatively traditional Fukushima prefecture. She is happy when she thinks about day-to-day life, but when her mind turns to the long term, as it inevitably does, she feels depressed. "This is temporary," she says. "We leave our house in the morning and we come home and it's temporary. It's like floating in the air."

Other people they know are struggling even more. Many of their current neighbours are out of work and stay at home all day. Some of Kenichi's former colleagues sent their wives and

A woman prepares lunch in her little partitioned unit at an evacuation centre, Miyako April 2011.

In the immediate aftermath of the nuclear accident, public-health experts worried about the possible risk from radiation. Subsequent analyses have shown that the prompt, if frantic, evacuation of areas around the reactors probably limited the public's exposure to a relatively safe level (see 'The evacuation zones'). But uncertainty, isolation and fears about radioactivity's invisible threat are jeopardizing the mental health of the 210,000 residents who fled from the nuclear disaster.

Researchers and clinicians are trying to assess and mitigate the problems, but it is unclear whether the Japanese government has the will, or the money, to provide the necessary support. Nor is it certain that the evacuees will accept any help, given their distrust of the government and their reluctance to discuss mental problems. This combination, researchers fear, could drive up rates of anxiety, substance abuse and depression.

The nuclear evacuees face a more difficult future than the survivors of the tsunami, which left nearly 20,000 dead or missing and caused billions of dollars in damage. "The tsunami-area people seem to be improving; they have more positive attitudes about the future," says Hirooki Yabe, a neuropsychiatrist at Fukushima Medical University, who has been working with both groups. Nuclear evacuees "are becoming more depressed day by day".

ESCAPE

Fukushima prefecture is a patchwork of orchards, rice paddies and fishing villages. In the 1970s and 80s, coastal residents welcomed nuclear power and two plants were built to supply electricity to Tokyo. Kenichi started working at Fukushima Daiichi in 1994, and at the time of the accident was a contract maintenance engineer. Yuka worked as a hospital nurse. The Togawas and their children, now aged 9, 12 and 15, lived in a four-room flat in Namie, a small, close-knit seaside town.

The family's life was upended at 14:46 on 11 March 2011. Kenichi was in the smoking room at the plant when he felt the ground shudder for several minutes. He ran back to his office, weaving around scattered desks and downed ceiling panels, and grabbed his driving licence and car keys. But he quickly found that traffic had clogged the routes out of the plant because the quake and tsunami had destroyed bridges and roads. Kenichi ditched his car and walked the remaining 8 kilometres home.

There he found that all his family members were safe, but he worried about the plant. In his job, he had overseen maintenance of the systems meant to cool the reactor in an emergency. If those had failed, he knew that a

meltdown would soon follow, spreading radiation to nearby towns. That night, as aftershocks rocked the house, his family slept fitfully with the lights and television on.

Kenichi was right to be alarmed. The tsunami had knocked out the generators that pumped cooling water into the reactor cores. As temperatures rose, the slender fuel rods full of uranium pellets began to warp. The meltdowns had begun.

Early the next morning, a siren wailed across Namie, signalling an evacuation. The Togawas were told to move to Tsushima, 30 kilometres to the northwest. After retrieving their car, the family set out, but the roads were choked with panicked residents and the Togawas ended up at a different evacuation centre. When Kenichi learned that the emergency diesel generators at the plant had also failed, he bundled the family into the car once more, hoping to reach Tsushima. "We have to run away," he remembers thinking in a panic.

On the way, Kenichi received a text message from a friend who worked at the power company's offices in Tokyo. Unit 1 had exploded, and radioactivity was spreading across Fukushima. The family drove from one full evacuation centre to the next, until they reached a dark, cramped gymnasium in Kawamata, around 40 kilometres northwest of the plant. There they were given a small patch of hardwood floor to call home. But they were still deeply worried about the radiation. "We didn't know much about radiation's effects, and we didn't know if Kawamata was safe or not," says Yuka.

Japan is used to natural disasters, and immediately after the tsunami hit, the country's

emergency services sprang into action. Groups of doctors and emergency workers from around the nation struck out for the northeast coast to begin search and rescue operations and to administer care. The medical university in Fukushima City became a hub. In the days and weeks after the accident, the university hospital took in seriously ill patients from the coast. It also found itself on the front line of the nuclear emergency: doctors used Geiger counters to screen evacuees' thyroid glands, which are particularly sensitive to radiation, and treated several workers from the nuclear plant, who had suffered high radiation exposures.

FIRST RESPONDERS

Mental-health experts were among the first responders, reflecting an ongoing change in Japan's attitudes towards mental health. For many years, Japan's modest but modern mental health services were geared to help only the most severely mentally ill. The society has traditionally paid little attention to more routine disorders such as depression. In recent years, however, the Japan Medical Association has started educating doctors about depression and suicide, and the national government has conducted public suicide-prevention campaigns.

Still, the quality of care remains patchy, and even before the accident, Fukushima prefecture was not a bright spot. Mental health was a not a priority for the rural, conservative region or its taciturn citizens. As a result, the tsunami and nuclear disaster strained the region's mental-health services to near breaking, says Yabe.

In the wake of the accident, most of the prefecture's resources were devoted to helping

EMPTY STREETS: In the aftermath of the nuclear disaster, 210,000 people had to evacuate Tomioka and other towns in Fukushima prefecture. Nearly 156,000 remain displaced from their homes.

those with established mental disorders. Yabe, for example, packed his car with antipsychotic and anticonvulsive medication and made runs to Soma City, where many evacuees had ended up. Mental-health professionals visited the cramped shelters elsewhere, but they tended to treat only the most severe cases of delirium and post-traumatic stress disorder.

The Togawas were among thousands of people left to their own devices by the overwhelmed doctors and counsellors. The family's first days in the cramped shelter are difficult to recall now, says Yuka, but what she can remember isn't pleasant: ill and elderly patients lying on the floor; ongoing fear about radiation; evacuees jumping queues and snatching food. "We were like dogs and cats without chains," she says.

With little guidance from the outside, the shelter's residents tried to organize themselves. Yuka volunteered her nursing skills, but after working for three days, she was filled with anger: why should she, a victim, have to spend all her time helping others, she wondered. Yuka locked herself in the family's car outside the shelter, "and just exploded and screamed and shouted and cried".

SUBTLE DAMAGE

As the evacuees struggled to adjust, so too did the doctors and psychologists at Fukushima Medical University. By May, the emergency response was mostly over and the hospital

had a new job — to assess the public's radiation dose. The task has proved tricky, says Shunichi Yamashita, a radiation health expert at Nagasaki University, who was brought in to head the Fukushima Health Management Survey. The radiation monitors around Daiichi were damaged or destroyed by the earthquake and tsunami, and the chaotic nature of the evacuation makes it difficult to assess how long and severely each person was exposed.

The few attempts made so far, however, have generally shown minimal risk. The health survey's latest assessment suggests that the dose for nearly all the evacuees was very low, with a maximum of only 25 millisieverts (mSv), well below the 100-mSv exposure that has been linked to an increased risk of cancer in survivors of the atomic bombs dropped on Hiroshima and Nagasaki in 1945. The World Health Organization also issued a reassuring report in May, saying that most evacuees from places like Namie received estimated doses between 10 and 50 mSv. It did note, however, that infants might have received a dose that could increase the risk of cancer in their still-developing thyroids.

Radiation specialists say that it is difficult to predict the health effects from such low doses. "I think it's likely that there will be increased cancer risks, but they will be very, very small," says Dale Preston, an independent statistician who has studied atomic-bomb survivors. "If you did a large study, I think your chance of

observing a statistically significant radiation-associated risk would be pretty low."

With that in mind, the health survey decided against following a fixed cohort to study the incidence of disease. Instead, it provides thyroid screening and other health checks to any evacuees who desire them. The hope is that the screenings themselves, along with the data collected, will help to reassure the public that the risks are low, says Yamashita.

Mental health has been a major component of the survey. In January 2012, researchers sent out questionnaires to all 210,000 evacuees to assess their stress and anxiety. The levels tabulated among the more than 91,000 respondents were "quite high", says Yuriko Suzuki, a psychiatrist at the National Institute of Mental Health in Tokyo. Roughly 15% of adults showed signs of extreme stress, five times the normal rate, and one in five showed signs of mental trauma — a rate similar to that in first responders to the attacks of 11 September 2001 in the United States. A survey of children, filled out by their parents, showed stress levels about double the Japanese average.

The stress has pushed some evacuees to breaking point. On a crisp day last November, Kenji Ookubo wandered through Iitate, a village 40 kilometres northwest of the plant, practising his golf swings in the empty streets. The town had been evacuated after the accident because it lay in the path of the plume of radiation blowing away from the plant. But Ookubo couldn't stand the temporary housing, where he had started drinking and suffered from stomach aches. After renting a room in Kawamata, he began squatting in his parents' abandoned home. "I came back just to run away from the stress," he says. With no job, and no prospects, "I can't see the future," he says.

It is a pattern seen frequently after major catastrophes, says Ronald Kessler, a professor of health-care policy at Harvard Medical School in Boston, Massachusetts. "In the short term, people get energized," he says. But when extensive damage or health problems prevent them from getting back to their old lives, depression and anxiety set in. "When something this big happens, it's just ridiculously daunting," he says. "At a certain point you just get worn down."

His own surveys of people evacuated after Hurricane Katrina, which struck the United States in 2005, show[1] that property loss and health concerns were the main causes of anxiety. Whereas many survivors of the Japanese tsunami have seen their homes rebuilt and lives restored, nuclear refugees are still dealing with both of those problems. Above all, the fear of radioactivity takes a unique toll. "It's something you don't feel; you don't notice what happened, and yet you understand that there are these long-term risks," says Preston. "It's scary."

Little is known about the long-term effects of
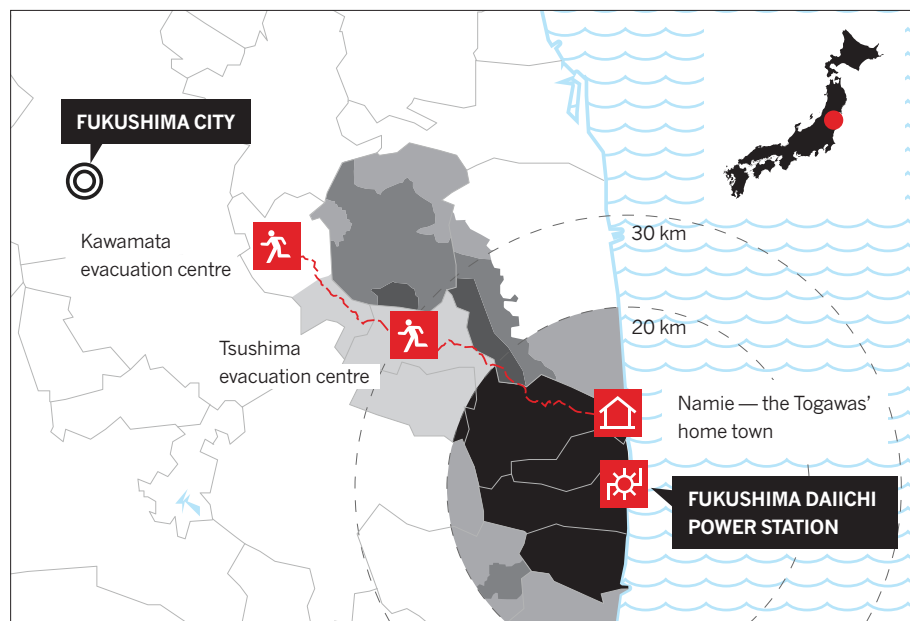
THE EVACUATION ZONES: During the first stages of the nuclear crisis at the Fukushima Daiichi power station, Japan evacuated a 20-kilometre radius around the plant, forcing the Togawa family to flee. They stopped at several centres before reaching Kawamata, where they currently reside. Japan later updated and expanded the evacuation zones.

- ■ Fully evacuated zone.
- ■ Dose rate of 20–50 millisieverts per year. Entry is restricted.
- ■ Dose rate >20 millisieverts per year. Some access allowed. No protective equipment required.
- ■ Residents may move freely but not stay over night.
- □ Planned evacuation area.

that fear, in part because nuclear accidents are so rare. But the 1986 disaster at the Chernobyl nuclear power plant in Ukraine suggests that fear of radiation can cause lasting psychological harm. Two decades after the accident, those who had evacuated as children complained of physical ailments more often than their peers, even though there was no difference in health[2]. And the mothers of those children suffered from post-traumatic stress disorder at about twice the rate of the general population, says Evelyn Bromet, a psychiatrist at the State University of New York in Stony Brook. Other studies of Chernobyl's aftermath found that evacuees had elevated rates of depression[3] and that a subset of clean-up workers committed suicide at a rate about 1.5 times that of the general population[4].

For Fukushima evacuees, says Bromet, "There's going to be a tremendous amount of health-related anxiety and it's not going to go away easily."

FEAR FACTOR

Yabe says that "radiophobia" remains a major problem among the Japanese refugees. A poll published last year by the Pew Research Center in Washington DC, for example, found that 76% of Japanese people believed that food from Fukushima was not safe, despite government and scientific assurances to the contrary. And many do not trust the government health surveys that found very few cases of significant radiation exposure among evacuees.

Yuka shares some of those concerns. She and Kenichi have educated themselves, and they have gained some reassurance from regular health checks and thyroid screenings. The children carry dosimeters provided by the health survey to collect radiation data and to calm public concerns. But Yuka wonders whether they will one day develop cancer.

At the moment, however, the family is preoccupied with practical concerns. The government has said that the Togawas can remain in their small flat until August 2014, but after that, Yuka says, they don't know what will happen. "The government officials say that they are working on it and that they are trying to construct public housing for those people who had to evacuate. But where? Nothing is clear." Whenever she and Kenichi think about the long term, they start to feel depressed.

The scientists involved with the Fukushima Health Management Survey have assigned a team of psychiatrists and nurses to make follow-up phone calls to individuals who had high scores for distress on the mental-health questionnaire. But only about 40% of adults responded to the questionnaires, and the researchers suspect that the most severely affected people did not participate. Even when the psychiatrists can connect, the evacuees usually don't stay on the phone for more than five to ten minutes. "Northern people are a very closed people, they

MAI NISHIYAMA
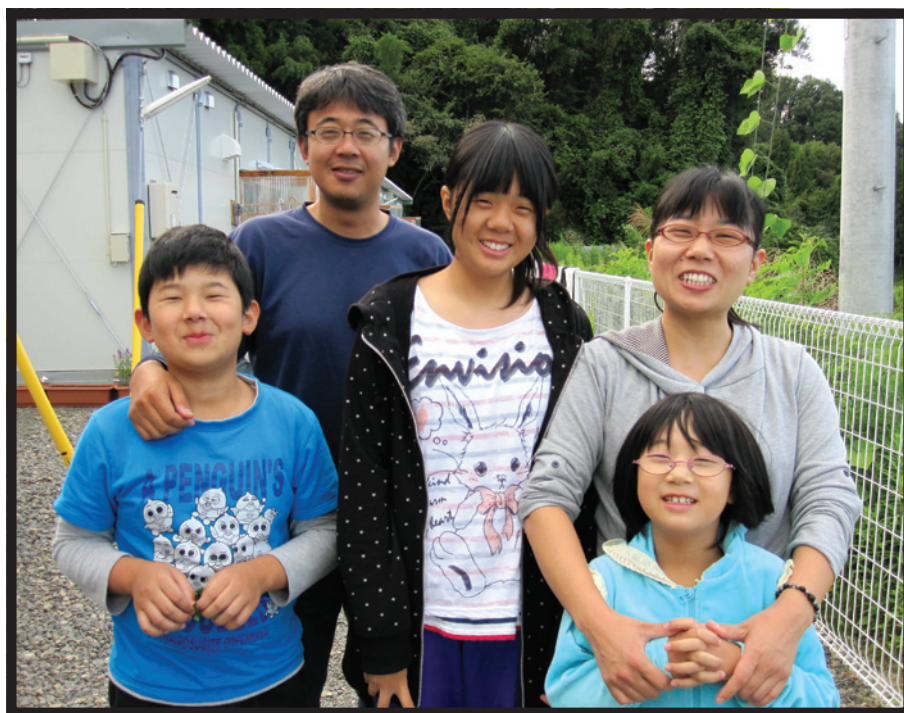
THE TOGAWAS: After nearly two years in temporary housing, the family worries about the future but is happy to be together. From left, Shoichiro, Kenichi, Rina, Yuka (back) and Kae (front).

don't really talk about their personal things, especially to somebody they've never met before," says Yabe.

Even when the psychiatrists identify problems, it is unclear what to do about them. Most evacuees, like the Togawas, are suffering from sub-clinical problems — mental anxiety and stress that affects their everyday life but does not require hospitalization or extensive therapy. There is no established treatment regime for such survivors from large disasters, says Suzuki.

Yabe suggests that walk-in clinics specializing in mental health could be set up throughout Fukushima prefecture to engage communities and help families. Suzuki says that involving large segments of the population in group-therapy sessions might be the way forward. Many say that it would help for evacuees to develop a sense of community — but the government has not fostered that. Temporary houses are "strung out like a railroad", says Bromet. The government could have built them "in a circle with a playground in the middle, or some obvious place for people to meet, but they didn't", she says.

Kessler says that unlike the tsunami survivors, whose grief will lessen over time, the nuclear evacuees could experience growing anxiety, particularly about radiation. "When everything has settled down, that will be a huge, rife issue," he predicts. Now is the best time to try to get ahead of these problems, he says. "There's a window of opportunity."

But the health survey lacks the funding for a more ambitious programme. The national government has given it just ¥3 billion (US$34 million) a year, but it is currently consuming about twice that amount, so the survey is under enormous financial pressure, says Seiji Yasumura, one of its leaders and an epidemiologist at Fukushima Medical University. So far, only 100 of the 210,000 evacuees have been interviewed face-to-face by mental-health experts.

Little by little, things are getting better for the Togawas. The children seem happy in their new school, and in September 2011 Kenichi found a job with the local government, clearing contaminated soil from the homes of neighbours. "He's worked so much overtime that his company is saying he has to have a break," boasts Yuka. She has found part-time work as a nurse in a local clinic. Her occasional outbursts sometimes cause tension with co-workers, but she enjoys speaking her mind: "I say what I want to say."

After filling out one of the health-survey questionnaires last year, Yuka got a flyer in the post inviting her to talk to someone over the phone. She thought about it but decided not to. "I don't feel like phoning. It's been nearly two years," she says. "I don't know what to say." ■ **SEE EDITORIAL P. 271**

Geoff Brumfiel *is a senior reporter at* Nature.

1. Galea, S. *et al. Arch. Gen. Psychiatr.* **64,** 1427–1434 (2007).
2. Bromet, E. J. *et al. BMC Publ. Health* **9,** 417 (2009).
3. Havenaar, J. M. *et al. Am. J. Psychiatr.* **154,** 1605–1607 (1997).
4. Rahu, K., Rahu, M., Tekkel, M. & Bromet, E. *Ann. Epidemiol.* **16,** 917–919 (2006).

COMMENT

YI LU/VIEWSTOCK/CORBIS

A hyena surveys a flock of flamingos in South Africa.

Time to model all life on Earth

To help transform our understanding of the biosphere, ecologists — like climate scientists — should simulate whole ecosystems, argue **Drew Purves** and colleagues.

No report from the Intergovernmental Panel on Climate Change would fail to mention global climate models. Yet the international bodies that are charged with addressing global challenges in conservation — including the Intergovernmental Platform on Biodiversity and Ecosystem Services, which holds its first plenary meeting next week in Bonn, Germany — cannot refer to analogous models of the world's ecosystems. Why? Because ecologists have not yet built them.

General circulation models, which simulate the physics and chemistry of Earth's land, ocean and atmosphere, embody scientists' best understanding of how the climate system works and are crucial to making predictions and shaping policies. We think that analogous general ecosystem models (GEMs) could radically improve understanding of the biosphere and inform policy decisions about biodiversity and conservation. Currently, decisions in conservation are based on disparate correlational studies, such as those showing that the diversity of bird species tends to decline in deforested landscapes. GEMs could provide a way to base conservation policy on an understanding of how ecosystems actually work.

Such models could capture the broad-scale structure and function of any ecosystem in the world by simulating processes — including feeding, reproduction and death — that drive the distribution and abundance of organisms within that ecosystem. Ecologists could apply a GEM to African savannas, for instance, to model the total biomass of all the plants, the grazers that feed on the plants, the carnivores that feed on the grazers and so on. Over time, the flows of energy and nutrients could be mapped between them. All of the organisms would be grouped not by species, but according to a few key traits such as ▶

whether they are plants, birds or mammals, cold blooded or warm blooded, diurnal or nocturnal. By encoding processes such as migration and predation into simple mathematical and computational forms, ecologists could model what happens to the various groups over time.

Metrics such as the diversity of animal types inhabiting the grasslands could be used to assess the savannas' health, stability and resilience, and to analyse the fate of particular groups of organisms such as top predators. Ecologists could explore how these attributes might change in response to, say, climate change, the introduction of invasive species or poaching. And, because the rules of play are likely to be broadly similar no matter what the ecosystem, the GEM could equally be applied to forests, lakes or the remotest parts of the ocean, providing a common framework for understanding and managing disparate ecosystems on local and global scales.

There are huge challenges to building GEMs — not least, obtaining the appropriate types of data to validate the models' predictions. But the difficulties are not insurmountable. Theories abound for describing the processes that drive ecosystems, many of which are backed up by data.

BUILDING A PROTOTYPE

Over the past two years, we at Microsoft Research and at the United Nations Environment Programme World Conservation Monitoring Centre, both in Cambridge, UK, have built a prototype GEM for terrestrial and marine ecosystems. Called the Madingley model, it uses real data on carbon flows as a starting point. We have hit all sorts of computational and technical hurdles, and are expecting more as we develop the model. Yet the project demonstrates that building GEMs is possible. From the relationship between the mass of individual organisms and how long they live, or the effects of human perturbations such as hunting, to the distribution of

Modelling every organism in an ecosystem such as a tropical rainforest would be impossible.

biomass across Earth (see 'Model life'), the model's outputs are broadly consistent with current understanding of ecosystems.

Modelling ecological phenomena at various scales is not new. Conservationists frequently use models to predict how much habitat fragmentation an endangered species can tolerate. But few ecologists have tried to build models of how general ecosystem properties emerge from the interactions of individuals. One attempt — known as Ecopath with Ecosim[1] — is being developed at the University of British Columbia's Fisheries Centre to address management and ecological questions. This combines modelling at the phytoplankton scale with that at the level of fisheries and marine mammals. Another, called Atlantis[2], developed by Australia's Commonwealth Scientific and Industrial Research Organisation, incorporates biophysical, economic and social factors to provide an integrated tool for modelling marine ecosystems. But these consider fewer processes and organisms than would a GEM.

Throughout the history of ecology, most researchers have resisted abstraction because

ecological complexity is so obvious in nature. Ignoring the myriad shapes and colours of different bird species, for example, seems instinctively wrong. Instead, ecologists have tended to stress the importance of species identification as well as a vast number of ecological processes, from individual adaptation to the social dynamics of groups. Many in the field also emphasize that findings in one ecosystem do not generalize to others, and that randomness and history could be as important in affecting some particular measurement as any deterministic rules[3]. But comprehensive species-specific data will always be in short supply (at least 80% of the millions of species on Earth are undescribed[4]), so a better understanding of ecosystems demands a broad-brush approach.

Building a GEM will require different types of data — to help define the ecological rules at play, to provide reasonable starting conditions for the simulation (such as a realistic ratio of herbivores to top predators) and to evaluate the model's predictions. For these three goals, enough data are available to get started, although information on ecological processes far outweighs the rest.

Metabolic rates, for instance, have been measured in hundreds of animals in the lab, and researchers in the field have documented lifespans, growth rates and reproductive success for thousands (in some cases, millions) of birds, mammals, plants and bacteria. Ecologists have also mathematically determined numerous 'rules of existence' for some organisms, such as that an animal's metabolic rate is proportional to its mass raised to a power of around 0.70 (ref. 5).
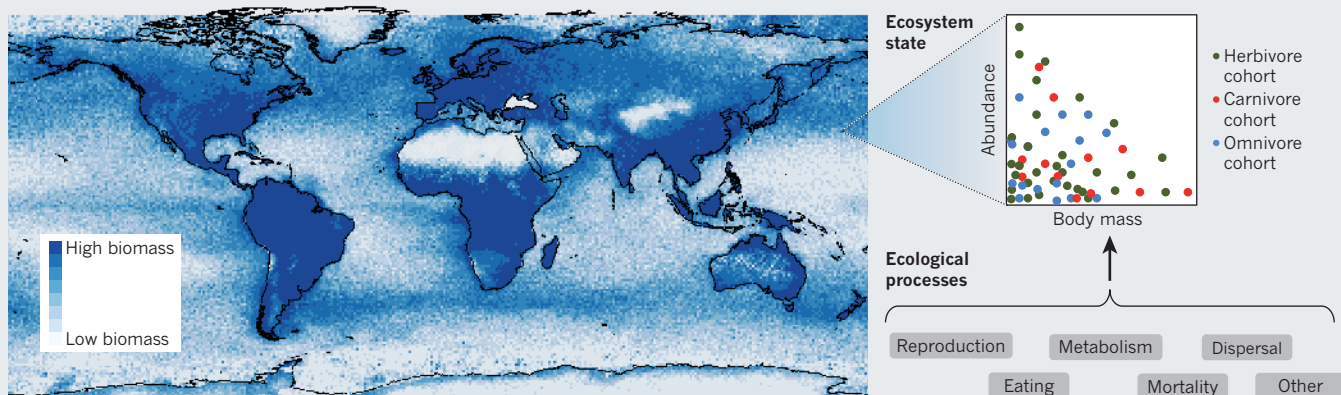
MODELLING BEHAVIOUR

Obviously, modelling every organism within an ecosystem is impossible. (We estimate that it would take a standard laptop computer around 47 billion years to model for 100 years every multicellular animal within just one of the 1-degree grid cells covering

SIMON LONG/FLICKR/GETTY

MODEL LIFE

Variation in biomass across the world simulated by the Madingley model for terrestrial and marine ecosystems. Fundamental ecological processes, encoded into simple computational forms, determine the abundance and body mass of organisms (grouped into cohorts for simplicity) and so indicate the state of ecosystems.

High biomass

Low biomass

Ecosystem state

Abundance

Body mass

- Herbivore cohort
- Carnivore cohort
- Omnivore cohort

Ecological processes

Reproduction   Metabolism   Dispersal

Eating   Mortality   Other

From hunting zebra to filter feeding, the process of predation in all ecosystems plays by similar rules.

Earth.) Yet certain computational techniques have been developed, mainly in marine ecology, that could allow researchers to model entire ecosystems using rules about the behaviour of individuals.

One approach is to model collections of organisms or 'cohorts'. The idea is that within a cohort, individuals are similar enough to be considered identical. For a shoal of small herbivorous fish meeting a cloud of plankton, say, ecologists could calculate the feeding rate for an exemplar fish and then apply that rate to the whole shoal. (In the simulation just described, we found that grouping organisms into cohorts on the basis of body size, functional group such as omnivore or carnivore and a few other traits reduced the computation time to 10 hours).

The biggest stumbling block to constructing GEMs (after convincing ecologists that they can and should be built!) is obtaining the data to parameterize and validate them.

Records of what species of plants and animals live in the world's forests, grasslands and oceans are often available to some extent, but far fewer data exist on the abundance of those species. And almost no data have been collected on the properties of whole ecosystems, such as on the distribution of body sizes from plankton to whales. Marine trawl surveys carried out for research or to assess fish stocks probably come closest to providing this type of information, although even these are restricted as to what size range of organisms they survey.

A new programme of data gathering is easy to envisage. Using automated cameras and image recognition, it should be possible to sample thousands of animals and determine their approximate size and what broad group they belong to: reptile or mammal, flying or non-flying. Motion-activated cameras used by conservationists and wildlife enthusiasts already produce tens of thousands of images of fish, birds and mammals every day. And stored away in numerous research institutions are vast samples of insects collected in

traps that suck them out of the air, and data from continuous plankton recorders towed beneath ships for millions of kilometres.

Naturally, a major new data-gathering programme would be costly. But globally, governments already spend billions of dollars on satellite observations of vegetation and habitat distribution, fisheries surveys, forest inventories and species surveys. Diverting a small fraction of these funds to gathering the data needed to develop and evaluate GEMs could pay dividends. A first step would be for governments around the world to support programmes similar to the National Ecological Observatory Network — an international cooperation funded by the US National Science Foundation to manage large-scale collection of ecological and climate data.

> *"Ecological systems do not have the equivalent of the precise laws used by climate scientists."*

To reduce costs and to harness the power of citizen science, data collection could even be crowd-sourced. Rapidly growing websites such as iNaturalist and eBird (on which users can share their observations of wildlife) currently focus on traditional species identification. Such sites could potentially collate an extraordinary amount of information on functional groups of organisms and traits such as body size.

TRUSTED ADVICE

Constructing realistic GEMs is one thing. The real challenge is to produce models from which the predictions are trustworthy enough to guide the decisions of conservationists and policy-makers. Recent progress in computational statistical methods offers a way for ecologists to formally build trustworthiness into models. For instance, tools are available to quantify the uncertainty associated with models' predictions. A healthy crop of alternative, competing GEMs will be crucial, together with mechanisms that enable their

fair assessment. In blind-testing, for example, different models could be used to predict an ecosystem property that has been measured but not reported, allowing the models to be ranked in terms of how well they do.

We are not proposing that GEM predictions (which will always be simplistic) provide the only guide to conservation policy and the management of ecosystems. But coupled with models from other fields, such as economics and epidemiology, they could offer a means of managing human actions and the biosphere in an integrated, consistent and evidence-based way. Far from eclipsing traditional ecological research, GEMs would draw on it and give such work more focus. Using GEMs, ecologists could identify processes that are poorly understood yet crucial to ecosystem structure and function, rather than delve deeper into well-studied areas.

Ecological systems do not have the equivalent of the precise laws used by climate scientists. This is a significant challenge to building GEMs, along with the complexity of nature, the small number of GEM-like models under development and the paucity of data with which to constrain them. But just by attempting to build general models, ecologists will find out what they need to know to truly understand ecosystems. ∎

Drew Purves *is head of the Computational Ecology and Environmental Science Group at Microsoft Research in Cambridge, UK.* **Jorn Scharlemann, Mike Harfoot, Tim Newbold, Derek P. Tittensor, Jon Hutton, Stephen Emmott.**
e-mail: dpurves@microsoft.com

1. Christensen, V. & Walters, C. J. *Eco. Mod.* **172,** 109–139 (2004).
2. Fulton, E. A. *et al. Fish Fish.* **12,** 171–188 (2011).
3. Boyd, I. L. *Science* **337,** 306–307 (2012).
4. Mora, C., Tittensor, D. P., Adl, S., Simpson, G. B. & Worm, B. *PLoS Biol.* **9,** e1001127 (2011).
5. Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. & West, G. B. *Ecology* **85,** 1771–1789 (2004).

Full author affiliations accompany this article online at go.nature.com/ob2a2p.

Two years of blogging the NIH

Sally Rockey, deputy director for extramural research at the US National Institutes of Health, reflects on the second anniversary of her precedent–setting blog.

In 2010, a few of my staff members began pressuring me to start a blog. Although my friends and colleagues will attest that I'm not a shy or soft-spoken individual, I was, to put it mildly, reticent. Blogging was relatively new territory for government officials at the time. There were some outstanding examples of blogs by science-oriented government offices, such as the White House Office of Science and Technology Policy, which by then had been running for a couple of years (www.whitehouse.gov/administration/eop/ostp/blog). Jeremy Berg, former director of the National Institute of General Medical Sciences, had also successfully established a following on the blog Feedback Loop (https://loop.nigms.nih.gov), so I knew that it could be done well. But how would people respond to a National Institutes of Health (NIH) blog on extramural research policy? And how would it fit in to our work of supporting biomedical research?

My staff knew that I wanted transparency to be a centrepiece of my tenure at the NIH, and felt that the biomedical research community would find it refreshing to hear more from NIH management. Revealing how we do our business as we are developing policies and processes, and presenting data and information that shed light on the impact of our decisions, seemed like the right thing to do.

On 19 January 2011, I launched the Rock Talk blog (http://nexus.od.nih.gov/all/rocktalk). Two years later, I'm delighted that I did (even though I have taken heat for using too many exclamation marks!). As of 2 January 2013, I've written 107 blog posts, averaging about 40,000 page views a month. A steady stream of comments keeps me on my toes (see 'Top ten hits'). For the first time, the NIH was presenting data from all of its institutes and centres, and was sharing its analysis of funding trends on a forum where scientists could give their own opinions and reactions directly.

My first data-sharing post looked at trends in funding at various types of research institution — showing, for instance, that the majority of NIH funds go to medical schools, and that distributions to for-profit institutions have increased since the 1980s. Within two weeks, the post generated nearly 90 comments, about everything from the ethics of training non-US scientists to the importance of including more MD-trained investigators.

A SENSE OF COMMUNITY

We learned a lot about our constituents' interests and needs through the blog, and we have been able to highlight behind-the-scenes data, actively engage the community in policy-making and provide insight into our decisions. Without public input, effective and impactful policy cannot be created. My blogging experience has convinced me that using social-media platforms is one effective way for science-funding agencies

TOP TEN HITS

Ranked by number of individual page views, the ten most popular entries on the NIH Rock Talk blog generated varying numbers of comments.

Blog post title	Summary	Number of comments
1. Age Distribution of NIH Principal Investigators and Medical School Faculty	Compares the average age of NIH principal investigators (PIs) to that of medical-school faculty members, showing that more scientists are becoming PIs later in life and retiring later.	63
2. Paylines, Percentiles and Success Rates	Explains the relationship between the various scores applicants receive for grants, and how those correlate to success rates.	52
3. 2011 Success Rates, Applications, and Investigators	Statistics on funding for the 2011 fiscal year.	12
4. What's Behind the 2011 Success Rates?	Details on the 2011 success rates, and an explanation of how to find this information on RePORT.nih.gov.	46
5. How Do You Think We Should Manage Science in Fiscally Challenging Times?	Interactive data on the effect of potential policy changes, and an invitation for scientists to comment on how best to manage the NIH during times of austerity.	253
6. Does Your Degree Matter?	Funding rates for NIH applicants with PhDs, MDs or both.	19
7. The A2 Resubmission Policy Continues: A Closer Look at Recent Data	Data showing that preventing applicants from resubmitting proposals twice enables the NIH to award more proposals earlier in the application process, and doesn't penalize young investigators.	79
8. President's 2013 Budget and NIH Research Grants	Proposed 2013 budget and follow-up post to 'How Do You Think We Should Manage Science in Fiscally Challenging Times?'.	56
9. Postdoctoral Researchers — Facts, Trends, and Gaps	A closer look at postdoc support, scientists' salaries and time to first tenure-track job, using data from the US National Science Foundation's Survey of Earned Doctorates.	27
10. Piloting the $1.5M Special Review	Announcement of the decision to add an extra layer of review for well-funded applicants.	55

to successfully support research.

The blog is not the official vehicle for communicating policy changes. We have the NIH Guide for Grants and Contracts for that. And we have official channels for soliciting public feedback — through requests for information in the NIH Guide, for example. But the blog allows me to extend that conversation to people worldwide, many of whom I would not be able to reach in other ways. People are sometimes concerned that offering a dissenting opinion to NIH officials might affect their chances of getting funding. Although that is absolutely not the case, one advantage the blog does have over some other channels is that it allows people to remain anonymous if they wish.

Another attribute of the blog — which I think should be mirrored by most government blogs — is how we communicate current events in real time. For example, when Hurricane Sandy hit the US east coast in October 2012, we posted advice on what awardees should do if their labs were affected by the storm, and highlighted relevant resources and guidance for dealing with natural disasters.

I am often asked whether I really read and respond to comments in person. I do check them every day, on average; some I answer myself and some with the assistance of my blog team (made up of four staff who help regularly, plus five or so who contribute when needed). At times, the criticism of the NIH or even of me is harsh, and can feel personal. But I understand. In this time of tightening budgets, the topic of funding is very personal, and I understand the commenters' frustrations. The blog is moderated and we have had to screen out a very few comments that are off-topic, constitute a personal attack on another

commenter or discuss the details of individual grants or applications (see the commenting policy at go.nature.com/ipwxqs).

The blog post that has received the most comments to date, 'How Do You Think We Should Manage Science in Fiscally Challenging Times?', was also one of the shortest. In fewer than 300 words, it asked scientists for their thoughts on how best to manage the NIH during times of austerity. This level of transparency into our thinking about future funding strategies was a novelty. The post also linked to interactive data charts that allowed readers to manipulate variables to see how their proposed changes to the size and numbers of grants we fund would affect success rates. We received more than 250 comments, including suggestions to take a closer look at how much funding was being given to individual scientists. We reviewed the suggestions and instituted our new Special Council Review policy to provide additional consideration of applications from principal investigators with existing NIH research project grants of US$1 million or more in direct costs (see J. M. Berg *Nature* **489**, 203; 2012).

> *"Conversation is an integral part of the policy-making process."*

Blog posts containing data have clearly been the most popular (see table). Another common theme of my posts is debunking urban myths about NIH-supported research. Ideas for these posts come from blog readers' comments, grantee e-mails, discussions with attendees at research conferences and more. For example, in 'Myth Busting: Number of Grants per Investigator' we showed that the myth that many principal investigators

continuously add to their bounty by piling on more and more NIH awards is untrue. The data we provided showed that the investigators who received the top 20% of funding in 2009 averaged only 2.2 grants each. (This example also illustrates how readers interact with the blog — they noticed that we had posted the wrong information in one of our graphs, which led to a follow-up post correcting the information and elaborating on the data we presented.)

Blogging is a lot of work, but it is also fun and enlightening. It would not be possible without the help of my Rock Talk blog and data teams, who provide statistical analyses, help me to decide on which topics to cover and moderate and respond to comments.

I encourage others in science policy to go into blogging with their eyes open, recognizing the level of effort required to care for and maintain a blog. Potential bloggers should be aware of the long-term commitment needed to determine which issues are blog-worthy and of interest to the community. Rock Talk has sometimes covered topics that reveal NIH idiosyncrasies. But I think that is what makes the blog real and helps us to connect with the community.

If you are not a reader of Rock Talk, I hope you will check it out. If you are, hopefully you know by now that I welcome your ideas, through either commenting on the blog or even starting a blog of your own. Conversation is an integral part of the policy-making process, and helps the biomedical research enterprise to keep on rockin'. ■

Sally Rockey *is deputy director of extramural research at the National Institutes of Health, Bethesda, Maryland 20892, USA.*
e-mail: sally.rockey@nih.gov

Jaguars were scarce even before they came under threat from humans.

CONSERVATION

The commonness of rarity

Stuart Pimm enjoys an exploration of what we can learn from rare species, if we care enough to keep them.

No writer of fiction would dare to invent the great naturalist Alfred Russel Wallace's adventures. In 1848, he funded his trip to the Amazon by collecting butterflies. Returning years later, his ship burned and sank mid-Atlantic; he lost his specimens, survived, and then set out for Asia. From Sarawak, Wallace nailed evolution's key consequences: "Every species has come into existence coincident both in space and time with a pre-existing closely allied species." He did it in so few words, he could have published it in *Nature*. But the scientific world largely ignored him.

When Wallace later worked out what caused these laws — that is, these patterns in both time and space — he wrote to Charles Darwin for support. What islands in tropical Asia did for Wallace, the Galapagos had already done for Darwin. Familiar history follows. Less familiar is that Wallace and Darwin had discovered ecological rarity, the subject of Eric Dinerstein's engaging new book.

As Dinerstein explains in *The Kingdom of Rarities*, "many, many species on Earth are rare, but few people … are even aware of this". He gives us a kind of zoological travelogue, observing rare species across the planet and contemplating, as he does so, why rarity is profoundly important for our understanding of nature and our efforts to conserve it. Dinerstein — who is lead scientist of conservation science at wildlife charity WWF — starts by explaining that there are two kinds of rarity. Some species are

widespread but sparse. Others are found only in very special places.

Large predators, such as jaguars, are an example of the sparse, widespread species. Dinerstein tracks these animals in South America and it soon becomes obvious why they are rare. Large-bodied species need a lot of food — and that demands large areas to support enough prey. Large fierce predators figure prominently in our imagination ("Lions and tigers and bears, Oh my!" says Dorothy in *The Wizard of Oz*). Dinerstein explains why Dorothy's fears are increasingly unfounded. We hunt these creatures to extinction — all too easy, given that they are rare to start with. This sets further problems in train: their herbivorous prey often become superabundant and decimate vegetation on which yet other species depend.

Most of the book is, however, about rare species in special places, giving it a flavour of high adventure. We travel to a remote mountain top in the Indonesian province of Papua, accessible only by helicopter, through terrifying descents in thick mist. The species here are so rare that some are new to science. They live in isolation from "closely allied species" — and isolation from humans, too, as is clear from the abundance and tameness of large-bodied species in the area. There are species of marsupial tree-kangaroo, but no monkeys, leopards or tigers. As Wallace proposed, those species could walk to Bali and Borneo, but no farther, when sea levels were lower during glacial advances.

Dinerstein then takes us to mainland Southeast Asia, travelling with the first Western scientists to go there after the Vietnam War. Extensive bombing and the bush-meat trade make this an unlikely place for rare species to survive. Indeed, the rhinoceros (*Rhinoceros sondaicus*) is now extinct in Vietnam. In Hanoi, Dinerstein unexpectedly finds himself sitting next to General Vo Nguyen Giap, a "superb tactician" in the conflict with the United States. Now a conservationist, Giap says "the forest is our friend", adding that "it hid us and provided shelter". Dinerstein thinks that Vietnam has become "the last sanctuary for the region's forest-dwelling rarities": the forest hid them too.
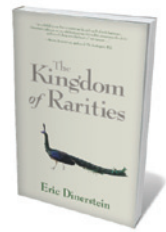
The Kingdom of Rarities
ERIC DINERSTEIN
Island: 2013.
312 pp. $29.95

MARK CARWARDINE/ARDEA.COM

STEVE GETTLE/MINDEN PICTURES/FLPA

Unexpectedly, Dinerstein then takes us to the United States, to Michigan, to visit the Kirtland's warbler (**pictured**), another rare species with a small range. Its existence suggests that the special places that host rare species can be anywhere, but then begs the question of where rare species are found most often. There are salient laws about diversity, beyond those that Wallace described.

First, most species have small ranges. Only a few species have large geographical ranges, but they are by far the most familiar to us. I've watched common kingfishers (*Alcedo atthis*) on British streams and on lakes in Japanese cities. To see other species in the same genus, however, I had to travel to Madagascar and Bali, and would need to travel to different islands in the Philippines to see two more.

Second, species with small ranges tend to be concentrated in special places — such as coastal Brazil or the eastern Himalayas. Most are tropical. Moreover, these areas are generally not where the greatest numbers of species are found in one place. So Dinerstein doesn't dwell on the Amazon: it has many species in any given location, but most have large ranges.

I sense that the Amazon disappointed Wallace too. It provided species, but not the unique ones his patrons craved. Darwin had no better luck in Patagonia, where the *Beagle* spent most of its famous voyage. When Wallace finally tripped over species with small, idiosyncratic ranges in the islands of Southeast Asia, and Darwin stopped briefly in the Galapagos on the way home, species rarity changed science forever.

Neither Darwin nor Wallace visited Hawaii, where Dinerstein's most poignant chapter is set. Different Hawaiian islands once had different "closely allied species". Dinerstein marvels at birds with strange names — the 'akiapola'au, for example — and even stranger beaks, and at lovely lobelias with flowers that fit other strange beaks. He sees a few of these in the wild; but most are extinct, just specimens in museums. Those that remain hang on in low numbers in the face of numerous threats such as habitat loss. When species such as these disappear forever, they take into oblivion all that we might learn about rarity and all that it tells us about the natural world. ∎

Stuart Pimm *is professor of conservation at the Nicholas School of the Environment, Duke University, Durham, North Carolina, USA, and author of* The World According to Pimm: a Scientist Audits the Earth.
e-mail: stuartpimm@me.com

STATISTICS

Number-crunching in the raw

David Vaux welcomes a primer on how statistics mesh with life, from salesmen's spin to political polls.

Ever wanted to encapsulate a quarter-back's performance in a single number, wondered whether a bowling score was a descriptive statistic, or wished for a mathematical dissection of the Monty Hall box-choice problem? Then this is the book for you.

Like St Jude, Charles Wheelan has an interest in saving lost causes, and he has been invoking nudity to help in his quest. In an earlier book, *Naked Economics* (W. W. Norton, 2002), he attempted to bring some joy to the 'dismal science'. With *Naked Statistics: Stripping the Dread from the Data*, he faces another tough challenge: after all, statisticians have been described as people who are good with numbers but lack the personality to be accountants.

But this book is not intended for statisticians, because it doesn't go into the theory or mathematics in depth. Nor is there enough detail for clinicians, researchers or students. Instead, *Naked Statistics* targets the layperson, especially North American sports fanatics who want a deeper understanding of baseball, football, basketball, bowling, running and shooting statistics. But there is still plenty of thought-provoking material for the rest of us — even an amateur statistician like me — on stat-rich areas from politics, commerce and polling to the stock market.

There are, for instance, chapters explaining the spins used by salesmen, advertisers and politicians. One example is the blind taste test set up by a US beer company in the early 1980s, involving people loyal to another brand. Roughly half chose the tester's beer, but as Wheelan notes, the similar taste of mid-range US beers could mean that the result was down to the law of probabilities. Wheelan is excellent on political polls, margins of error, and the importance of questioning a truly representative sample of the population, which is why surveys of sexual practices and internet polls can generate spurious results.

There is sound advice on what Wheelan reveals as different types of modern gambling. He explains how in the stock market you cannot expect to do better than the market index; that the only insurance you should buy is when you can't afford to easily cover the loss yourself; and that lottery tickets are for chumps. I would have liked more on the historical role of gambling, and how it was used by Fermat and Pascal to establish the mathematical basis of probability and statistics. It would also have been welcome to have practical tips other than "Don't do it", such as the theory and practicality of counting cards at casinos, and how to compare and judge risk.
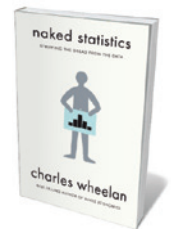
Naked Statistics also covers descriptive statistics (which allows you to describe how a population is distributed), inferential statistics (which can be used to compare populations or answer hypotheses), and correlation and regression analysis (which can be used to see whether data, such as diet and health, are related). Wheelan explains their uses and key shortcomings, but without formulae and problems to solve. The appendices to some of the chapters do include more mathematical detail, and will give you a deeper understanding, especially if you study them in conjunction with the superb, free, statistical videos on the Khan Academy website (http://www.khanacademy.org).

For a deeper explanation of statistics and a glimpse at a future in which we are not subject to the tyranny of *p* values — the most misunderstood statistical concept in research — I recommend Geoff Cumming's *Understanding the New Statistics* (Routledge, 2011).

And if you would like to know more about the areas Wheelan points to as hotbeds of use and abuse, take a look at *Proofiness* by Charles Seife (Viking, 2010) for the political angle, and *Bad Pharma* by Ben Goldacre (Fourth Estate, 2012), which looks at medical research and the pharmaceutical industry. These provide enough factual examples to get your blood boiling — without taking your clothes off. ∎

Naked Statistics: Stripping the Dread from the Data
CHARLES WHEELAN
W. W. Norton: 2013. 302 pp. $26.95

David L. Vaux *is deputy director of the Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia.*
e-mail: vaux@wehi.edu.au

Unexpectedly, Dinerstein then takes us to the United States, to Michigan, to visit the Kirtland's warbler (**pictured**), another rare species with a small range. Its existence suggests that the special places that host rare species can be anywhere, but then begs the question of where rare species are found most often. There are salient laws about diversity, beyond those that Wallace described.

First, most species have small ranges. Only a few species have large geographical ranges, but they are by far the most familiar to us. I've watched common kingfishers (*Alcedo atthis*) on British streams and on lakes in Japanese cities. To see other species in the same genus, however, I had to travel to Madagascar and Bali, and would need to travel to different islands in the Philippines to see two more.

Second, species with small ranges tend to be concentrated in special places — such as coastal Brazil or the eastern Himalayas. Most are tropical. Moreover, these areas are generally not where the greatest numbers of species are found in one place. So Dinerstein doesn't dwell on the Amazon: it has many species in any given location, but most have large ranges.

I sense that the Amazon disappointed Wallace too. It provided species, but not the unique ones his patrons craved. Darwin had no better luck in Patagonia, where the *Beagle* spent most of its famous voyage. When Wallace finally tripped over species with small, idiosyncratic ranges in the islands of Southeast Asia, and Darwin stopped briefly in the Galapagos on the way home, species rarity changed science forever.

Neither Darwin nor Wallace visited Hawaii, where Dinerstein's most poignant chapter is set. Different Hawaiian islands once had different "closely allied species". Dinerstein marvels at birds with strange names — the 'akiapola'au, for example — and even stranger beaks, and at lovely lobelias with flowers that fit other strange beaks. He sees a few of these in the wild; but most are extinct, just specimens in museums. Those that remain hang on in low numbers in the face of numerous threats such as habitat loss. When species such as these disappear forever, they take into oblivion all that we might learn about rarity and all that it tells us about the natural world. ∎

Stuart Pimm *is professor of conservation at the Nicholas School of the Environment, Duke University, Durham, North Carolina, USA, and author of* The World According to Pimm: a Scientist Audits the Earth.
e-mail: stuartpimm@me.com

STATISTICS

Number-crunching in the raw

David Vaux welcomes a primer on how statistics mesh with life, from salesmen's spin to political polls.

Ever wanted to encapsulate a quarterback's performance in a single number, wondered whether a bowling score was a descriptive statistic, or wished for a mathematical dissection of the Monty Hall box-choice problem? Then this is the book for you.

Like St Jude, Charles Wheelan has an interest in saving lost causes, and he has been invoking nudity to help in his quest. In an earlier book, *Naked Economics* (W. W. Norton, 2002), he attempted to bring some joy to the 'dismal science'. With *Naked Statistics: Stripping the Dread from the Data*, he faces another tough challenge: after all, statisticians have been described as people who are good with numbers but lack the personality to be accountants.

But this book is not intended for statisticians, because it doesn't go into the theory or mathematics in depth. Nor is there enough detail for clinicians, researchers or students. Instead, *Naked Statistics* targets the layperson, especially North American sports fanatics who want a deeper understanding of baseball, football, basketball, bowling, running and shooting statistics. But there is still plenty of thought-provoking material for the rest of us — even an amateur statistician like me — on stat-rich areas from politics, commerce and polling to the stock market.

There are, for instance, chapters explaining the spins used by salesmen, advertisers and politicians. One example is the blind taste test set up by a US beer company in the early 1980s, involving people loyal to another brand. Roughly half chose the tester's beer, but as Wheelan notes, the similar taste of midrange US beers could mean that the result was down to the law of probabilities. Wheelan is excellent on political polls, margins of error, and the importance of questioning a truly representative sample of the population, which is why surveys of sexual practices and internet polls can generate spurious results.

There is sound advice on what Wheelan reveals as different types of modern gambling. He explains how in the stock market you cannot expect to do better than the market index; that the only insurance you should buy is when you can't afford to easily cover the loss yourself; and that lottery tickets are for chumps. I would have liked more on the historical role of gambling, and how it was used by Fermat and Pascal to establish the mathematical basis of probability and statistics. It would also have been welcome to have practical tips other than "Don't do it", such as the theory and practicality of counting cards at casinos, and how to compare and judge risk.
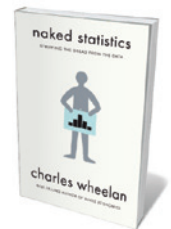
Naked Statistics also covers descriptive statistics (which allows you to describe how a population is distributed), inferential statistics (which can be used to compare populations or answer hypotheses), and correlation and regression analysis (which can be used to see whether data, such as diet and health, are related). Wheelan explains their uses and key shortcomings, but without formulae and problems to solve. The appendices to some of the chapters do include more mathematical detail, and will give you a deeper understanding, especially if you study them in conjunction with the superb, free, statistical videos on the Khan Academy website (http://www.khanacademy.org).

For a deeper explanation of statistics and a glimpse at a future in which we are not subject to the tyranny of *p* values — the most misunderstood statistical concept in research — I recommend Geoff Cumming's *Understanding the New Statistics* (Routledge, 2011).

And if you would like to know more about the areas Wheelan points to as hotbeds of use and abuse, take a look at *Proofiness* by Charles Seife (Viking, 2010) for the political angle, and *Bad Pharma* by Ben Goldacre (Fourth Estate, 2012), which looks at medical research and the pharmaceutical industry. These provide enough factual examples to get your blood boiling — without taking your clothes off. ∎

Naked Statistics: Stripping the Dread from the Data
CHARLES WHEELAN
W. W. Norton: 2013. 302 pp. $26.95

David L. Vaux *is deputy director of the Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia.*
e-mail: vaux@wehi.edu.au

The film *Fire In The Blood* focuses on the campaign for AIDS drugs in Africa.

FILM

Walking the edge at Sundance

Society and technology collide at this year's leading independent film festival, finds **Jascha Hoffman**.

Technology and imagination are cinema's warp and weft. So it is hardly surprising that the film industry has long been preoccupied with the leading edge of technological research — from *Metropolis* to *Transformers* — and with the ethical and social ambiguities that surround it. A range of movies tackle those grey areas at this year's Sundance Film Festival in Utah.

"We relate to the technology that we're offered," says John Nein, a senior film programmer at the festival. "We have to contend with the social and political forces around that technology."

The festival was founded in 1981 to help independent directors reach a wider audience. This year's science-and-technology-tinged offerings might convey less outright science than in years past, but they probe big issues — from the 'new' nuclear power to Google, WikiLeaks and the 1970s innovators who brought computers into our homes and our pockets.

A prominent pair of feature films deal with rivalry and uncertainty at the dawn of the digital era. Joshua Stern's *jOBS*, a biopic of Apple founder Steve Jobs, stars Ashton Kutcher as the young titan designing software and devices that will usher in the age of personal computing. *Computer Chess*, a comedy-drama about programmers teaching computers to play chess in the 1980s, underlines the programmers' anxiety to prove that computers could hold their own against humans. That remains a live issue as we begin to carry on conversations with the computers in our phones and cars. "The big philosophical questions may seem charmingly nostalgic, but lord knows we haven't answered them yet," director Andrew Bujalski wrote when raising a modest budget for the film online.

Two documentaries tackle controversial efforts to spread information on the Internet — and the resulting backlash. Ben Lewis's *Google and the World Brain* charts the Internet giant's decade-long effort to scan the world's books into a giant digital archive, prompting legal action from authors and librarians. And from prolific documentary director Alex Gibney comes *We Steal Secrets*, which traces the evolution of WikiLeaks from a site founded in Sweden in 2006 by Australian former hacker Julian Assange to an international movement that aims to make secrecy obsolete with actions such as the release of 250,000 US diplomatic cables in 2010 and 2011.

Medicine features heavily in another brace of documentaries, both of which show researchers and advocates racing against the clock to devise and disseminate cures. Sean and Andrea Fine's *Life According to Sam* traces the efforts of husband-and-wife US doctors Scott Berns and Leslie Gordon to understand the rare premature-ageing disease progeria, after their son was diagnosed at age two. Their work over the ensuing decade has led to gene discovery, preliminary drug trials and potential insights into the ageing process. Africa at the turn of the millennium is the focus of Dylan Gray's *Fire in the Blood*, which follows the public-health activists, including South African Zackie Achmat, who pushed governments and drug companies to make AIDS drugs available to Africans.

Even the films with an environmental theme have a technological edge. *Pandora's Promise*, directed by Robert Stone, is a documentary about environmentalists such as the British journalist Mark Lynas, who, after opposing nuclear power, now support the use of new technologies such as small modular nuclear reactors to reduce reliance on fossil fuels. The eco-thriller *The East*, directed by Zal Batmanglij and starring Brit Marling and Ellen Page, features an activist group that attacks corporations using their own products, such as harmful chemicals.

Some of the festival's more conceptual works use technology to encourage viewers to empathize with other creatures. *Coral: Rekindling Venus*, is a full-dome planetarium show by Australian artist Lynette Wallworth. Deep-sea photography and music by Antony Hegarty (of the New-York-based group Antony and the Johnsons) among others highlight the threats to coral reefs. The short film *Primate Cinema: Apes as Family*, by artist Rachel Mayeri, shows the reactions of real chimpanzees to a television programme acted by humans in chimp costumes, inviting us to ponder how our fellow primates see us.

A panel on science at the movies, organized by the Alfred P. Sloan Foundation's Science and Film programme, promises a frank discussion of "the best and worst of cinema science" while honouring the "creative, imaginative, speculative and even adventurous" spirit of scientists and film-makers. Panellists, including director Darren Aronofsky and theoretical physicist Lisa Randall, will also select a winner for this year's Sloan Prize for films about technology and science. ■

Jascha Hoffman *writes for* Nature *and* The New York Times. *He lives in San Francisco.* e-mail: jascha@jaschahoffman.com

SUNDANCE FILM FESTIVAL

Correspondence

Rat reality show blurs quality control

The governance of science, including the whole system of quality assurance, depends on specialist access to resources and publication. A new and radical engagement of the public in reality and crowd-sourced science is calling this principle into question.

Following the furore surrounding cancer claims for herbicide-resistant transgenic maize (*Nature* **489,** 484; 2012; and *Nature* **492,** 12; 2012), Russian scientists intend to run a 'rat reality show' that will be broadcast over the Internet around the clock. This year-long, real-time feeding experiment aims to test the safety of genetically modified food products (see go.nature.com/qkvzqe). Actually witnessing the experiment, the team suggests, will allow the public to draw their own conclusions. If the scientists are unable to fund the show from orthodox sources, they plan to crowd-source public funding.

It would be easy to dismiss such 'reality' experiments as a stunt — as frivolity leading to demagoguery. But social media are increasingly influencing mainstream scientific communication and could stimulate a spread in reality science, blurring the demarcations on which the legitimacy and quality assurance of science traditionally depend.

Jerome R. Ravetz, Peter Healey, Steve Rayner *Institute for Science, Innovation and Society, University of Oxford, UK.*
steve.rayner@insis.ox.ac.uk

Cancer goal: deadline will not erode trust

We at the US National Breast Cancer Coalition (NBCC) disagree with your assessment of our 2020 goal of knowing how to prevent breast cancer and its associated deaths (*Nature* **491,** 637; 2012). Working towards that

goal will not erode public trust, as you suggest.

Nearly 500,000 women worldwide died of breast cancer last year, despite billions of dollars being invested in research. Many scientists believe that current funding systems favour 'safe' research over bold new ideas. As a result, progress is incremental, leading to slightly better treatments, surgical interventions and radiation regimes. These may provide some benefit, but bring no end to the disease itself.

Scientists and the NBCC need to work together to reorder priorities and change the conversation and culture of science. Trust is not lost when advocates call for a deadline and provide a blueprint for meeting it. But it is eroded when scientific infrastructure is unaccountable to the people intended to benefit from its output; when there is not enough emphasis on translating research discoveries to the clinic; and when published results cannot be replicated and marginal advances are over-hyped. Meanwhile, more and more people lose their lives.

Take a calculated risk with us. Let's reach for what might in fact be possible.

Frances Visco *National Breast Cancer Coalition, Washington DC, USA.*
fmvisco@ breastcancerdeadline2020.org

Cancer goal: vaccine is cause for optimism

As the originator of the project to develop a preventive breast-cancer vaccine, which is supported by the US National Breast Cancer Coalition's (NBCC) 2020 deadline, I question your dismissal of this new approach (*Nature* **491,** 637; 2012).

You base your pessimism on the genomic complexity of tumours and the length of time that clinical trials would take to test such a vaccine. But the genomic complexity apparent in mature tumours is not relevant for developing a preventive vaccine: the immune system needs only to be preactivated with antigens presented by the nascent tumour. Unstable RNA processing in a tumour is likely to be a richer source of these antigens than genomic mutations.

It could indeed take a decade or more to validate the treatment using today's clinical standards. But trials could be made shorter and less expensive if, for example, there were definitive ways to detect very early tumours.

You say that "discovery does not answer to deadlines", but accomplishment can. Given the alternatives, we should embrace this effort.

Stephen Albert Johnston

Arizona State University, Tempe, USA.
stephen.johnston@asu.edu

Tie carbon emissions to consumers

Global carbon dioxide emissions soared from 22.7 billion tonnes in 1990 to 33.9 billion tonnes last year, despite 20 years of attempted mitigation (*Nature* **491,** 656–658; 2012). The sizeable economic gaps between nations are largely responsible for the international deadlock in climate negotiations. A radically new approach is needed.

One solution would be to allocate common but differentiated responsibility for mitigating emissions to individual consumption activities, rather than to countries. Profligate consumers from both developed and emerging countries are the worst offenders for generating non-essential emissions. They should be held accountable for those emissions, irrespective of where they are generated.

Emissions produced after reallocating responsibility in this way could be quantified for nations, lifestyles or even every product consumed. This could be achieved by integrating the 'top-down' consumption-based accounting methods that are used to determine national total consumption emissions with 'bottom-up' carbon-footprint calculations based on analysis of products' life cycles (G. P. Peters *Curr. Opin. Environ. Sustain.* **2,** 245–250; 2010). A living standard, together with a per capita emission quota, would be defined so that people worldwide could meet their basic living requirements without mitigation costs.

Such standardized measures would allow new cap-and-trade policies and carbon-taxing mechanisms to run smoothly and effectively across different consumption groups at a global scale (see M. Grubb *Nature* **491,** 666–667; 2012).

Correspondence

Rat reality show blurs quality control

The governance of science, including the whole system of quality assurance, depends on specialist access to resources and publication. A new and radical engagement of the public in reality and crowd-sourced science is calling this principle into question.

Following the furore surrounding cancer claims for herbicide-resistant transgenic maize (*Nature* **489,** 484; 2012; and *Nature* **492,** 12; 2012), Russian scientists intend to run a 'rat reality show' that will be broadcast over the Internet around the clock. This year-long, real-time feeding experiment aims to test the safety of genetically modified food products (see go.nature.com/qkvzqe). Actually witnessing the experiment, the team suggests, will allow the public to draw their own conclusions. If the scientists are unable to fund the show from orthodox sources, they plan to crowd-source public funding.

It would be easy to dismiss such 'reality' experiments as a stunt — as frivolity leading to demagoguery. But social media are increasingly influencing mainstream scientific communication and could stimulate a spread in reality science, blurring the demarcations on which the legitimacy and quality assurance of science traditionally depend.

Jerome R. Ravetz, Peter Healey, Steve Rayner *Institute for Science, Innovation and Society, University of Oxford, UK.*
steve.rayner@insis.ox.ac.uk

Cancer goal: deadline will not erode trust

We at the US National Breast Cancer Coalition (NBCC) disagree with your assessment of our 2020 goal of knowing how to prevent breast cancer and its associated deaths (*Nature* **491,** 637; 2012). Working towards that

goal will not erode public trust, as you suggest.

Nearly 500,000 women worldwide died of breast cancer last year, despite billions of dollars being invested in research. Many scientists believe that current funding systems favour 'safe' research over bold new ideas. As a result, progress is incremental, leading to slightly better treatments, surgical interventions and radiation regimes. These may provide some benefit, but bring no end to the disease itself.

Scientists and the NBCC need to work together to reorder priorities and change the conversation and culture of science. Trust is not lost when advocates call for a deadline and provide a blueprint for meeting it. But it is eroded when scientific infrastructure is unaccountable to the people intended to benefit from its output; when there is not enough emphasis on translating research discoveries to the clinic; and when published results cannot be replicated and marginal advances are over-hyped. Meanwhile, more and more people lose their lives.

Take a calculated risk with us. Let's reach for what might in fact be possible.

Frances Visco *National Breast Cancer Coalition, Washington DC, USA.*
fmvisco@ breastcancerdeadline2020.org

Cancer goal: vaccine is cause for optimism

As the originator of the project to develop a preventive breast-cancer vaccine, which is supported by the US National Breast Cancer Coalition's (NBCC) 2020 deadline, I question your dismissal of this new approach (*Nature* **491,** 637; 2012).

You base your pessimism on the genomic complexity of tumours and the length of time that clinical trials would take to test such a vaccine. But the genomic complexity apparent in mature tumours is not relevant for developing a preventive vaccine: the immune system needs only to be preactivated with antigens presented by the nascent tumour. Unstable RNA processing in a tumour is likely to be a richer source of these antigens than genomic mutations.

It could indeed take a decade or more to validate the treatment using today's clinical standards. But trials could be made shorter and less expensive if, for example, there were definitive ways to detect very early tumours.

You say that "discovery does not answer to deadlines", but accomplishment can. Given the alternatives, we should embrace this effort.

Stephen Albert Johnston

Arizona State University, Tempe, USA.
stephen.johnston@asu.edu

Tie carbon emissions to consumers

Global carbon dioxide emissions soared from 22.7 billion tonnes in 1990 to 33.9 billion tonnes last year, despite 20 years of attempted mitigation (*Nature* **491,** 656–658; 2012). The sizeable economic gaps between nations are largely responsible for the international deadlock in climate negotiations. A radically new approach is needed.

One solution would be to allocate common but differentiated responsibility for mitigating emissions to individual consumption activities, rather than to countries. Profligate consumers from both developed and emerging countries are the worst offenders for generating non-essential emissions. They should be held accountable for those emissions, irrespective of where they are generated.

Emissions produced after reallocating responsibility in this way could be quantified for nations, lifestyles or even every product consumed. This could be achieved by integrating the 'top-down' consumption-based accounting methods that are used to determine national total consumption emissions with 'bottom-up' carbon-footprint calculations based on analysis of products' life cycles (G. P. Peters *Curr. Opin. Environ. Sustain.* **2,** 245–250; 2010). A living standard, together with a per capita emission quota, would be defined so that people worldwide could meet their basic living requirements without mitigation costs.

Such standardized measures would allow new cap-and-trade policies and carbon-taxing mechanisms to run smoothly and effectively across different consumption groups at a global scale (see M. Grubb *Nature* **491,** 666–667; 2012).

Correspondence

Rat reality show blurs quality control

The governance of science, including the whole system of quality assurance, depends on specialist access to resources and publication. A new and radical engagement of the public in reality and crowd-sourced science is calling this principle into question.

Following the furore surrounding cancer claims for herbicide-resistant transgenic maize (*Nature* **489,** 484; 2012; and *Nature* **492,** 12; 2012), Russian scientists intend to run a 'rat reality show' that will be broadcast over the Internet around the clock. This year-long, real-time feeding experiment aims to test the safety of genetically modified food products (see go.nature.com/qkvzqe). Actually witnessing the experiment, the team suggests, will allow the public to draw their own conclusions. If the scientists are unable to fund the show from orthodox sources, they plan to crowd-source public funding.

It would be easy to dismiss such 'reality' experiments as a stunt — as frivolity leading to demagoguery. But social media are increasingly influencing mainstream scientific communication and could stimulate a spread in reality science, blurring the demarcations on which the legitimacy and quality assurance of science traditionally depend.

Jerome R. Ravetz, Peter Healey, Steve Rayner *Institute for Science, Innovation and Society, University of Oxford, UK.*
steve.rayner@insis.ox.ac.uk

Cancer goal: deadline will not erode trust

We at the US National Breast Cancer Coalition (NBCC) disagree with your assessment of our 2020 goal of knowing how to prevent breast cancer and its associated deaths (*Nature* **491,** 637; 2012). Working towards that

goal will not erode public trust, as you suggest.

Nearly 500,000 women worldwide died of breast cancer last year, despite billions of dollars being invested in research. Many scientists believe that current funding systems favour 'safe' research over bold new ideas. As a result, progress is incremental, leading to slightly better treatments, surgical interventions and radiation regimes. These may provide some benefit, but bring no end to the disease itself.

Scientists and the NBCC need to work together to reorder priorities and change the conversation and culture of science. Trust is not lost when advocates call for a deadline and provide a blueprint for meeting it. But it is eroded when scientific infrastructure is unaccountable to the people intended to benefit from its output; when there is not enough emphasis on translating research discoveries to the clinic; and when published results cannot be replicated and marginal advances are over-hyped. Meanwhile, more and more people lose their lives.

Take a calculated risk with us. Let's reach for what might in fact be possible.

Frances Visco *National Breast Cancer Coalition, Washington DC, USA.*
fmvisco@ breastcancerdeadline2020.org

Cancer goal: vaccine is cause for optimism

As the originator of the project to develop a preventive breast-cancer vaccine, which is supported by the US National Breast Cancer Coalition's (NBCC) 2020 deadline, I question your dismissal of this new approach (*Nature* **491,** 637; 2012).

You base your pessimism on the genomic complexity of tumours and the length of time that clinical trials would take to test such a vaccine. But the genomic complexity apparent in mature tumours is not relevant for developing a preventive vaccine: the immune system needs only to be preactivated with antigens presented by the nascent tumour. Unstable RNA processing in a tumour is likely to be a richer source of these antigens than genomic mutations.

It could indeed take a decade or more to validate the treatment using today's clinical standards. But trials could be made shorter and less expensive if, for example, there were definitive ways to detect very early tumours.

You say that "discovery does not answer to deadlines", but accomplishment can. Given the alternatives, we should embrace this effort.

Stephen Albert Johnston

Arizona State University, Tempe, USA.
stephen.johnston@asu.edu

Tie carbon emissions to consumers

Global carbon dioxide emissions soared from 22.7 billion tonnes in 1990 to 33.9 billion tonnes last year, despite 20 years of attempted mitigation (*Nature* **491,** 656–658; 2012). The sizeable economic gaps between nations are largely responsible for the international deadlock in climate negotiations. A radically new approach is needed.

One solution would be to allocate common but differentiated responsibility for mitigating emissions to individual consumption activities, rather than to countries. Profligate consumers from both developed and emerging countries are the worst offenders for generating non-essential emissions. They should be held accountable for those emissions, irrespective of where they are generated.

Emissions produced after reallocating responsibility in this way could be quantified for nations, lifestyles or even every product consumed. This could be achieved by integrating the 'top-down' consumption-based accounting methods that are used to determine national total consumption emissions with 'bottom-up' carbon-footprint calculations based on analysis of products' life cycles (G. P. Peters *Curr. Opin. Environ. Sustain.* **2,** 245–250; 2010). A living standard, together with a per capita emission quota, would be defined so that people worldwide could meet their basic living requirements without mitigation costs.

Such standardized measures would allow new cap-and-trade policies and carbon-taxing mechanisms to run smoothly and effectively across different consumption groups at a global scale (see M. Grubb *Nature* **491,** 666–667; 2012).

Correspondence

Rat reality show blurs quality control

The governance of science, including the whole system of quality assurance, depends on specialist access to resources and publication. A new and radical engagement of the public in reality and crowd-sourced science is calling this principle into question.

Following the furore surrounding cancer claims for herbicide-resistant transgenic maize (*Nature* **489,** 484; 2012; and *Nature* **492,** 12; 2012), Russian scientists intend to run a 'rat reality show' that will be broadcast over the Internet around the clock. This year-long, real-time feeding experiment aims to test the safety of genetically modified food products (see go.nature.com/qkvzqe). Actually witnessing the experiment, the team suggests, will allow the public to draw their own conclusions. If the scientists are unable to fund the show from orthodox sources, they plan to crowd-source public funding.

It would be easy to dismiss such 'reality' experiments as a stunt — as frivolity leading to demagoguery. But social media are increasingly influencing mainstream scientific communication and could stimulate a spread in reality science, blurring the demarcations on which the legitimacy and quality assurance of science traditionally depend.

Jerome R. Ravetz, Peter Healey, Steve Rayner *Institute for Science, Innovation and Society, University of Oxford, UK.*
steve.rayner@insis.ox.ac.uk

Cancer goal: deadline will not erode trust

We at the US National Breast Cancer Coalition (NBCC) disagree with your assessment of our 2020 goal of knowing how to prevent breast cancer and its associated deaths (*Nature* **491,** 637; 2012). Working towards that

goal will not erode public trust, as you suggest.

Nearly 500,000 women worldwide died of breast cancer last year, despite billions of dollars being invested in research. Many scientists believe that current funding systems favour 'safe' research over bold new ideas. As a result, progress is incremental, leading to slightly better treatments, surgical interventions and radiation regimes. These may provide some benefit, but bring no end to the disease itself.

Scientists and the NBCC need to work together to reorder priorities and change the conversation and culture of science. Trust is not lost when advocates call for a deadline and provide a blueprint for meeting it. But it is eroded when scientific infrastructure is unaccountable to the people intended to benefit from its output; when there is not enough emphasis on translating research discoveries to the clinic; and when published results cannot be replicated and marginal advances are over-hyped. Meanwhile, more and more people lose their lives.

Take a calculated risk with us. Let's reach for what might in fact be possible.

Frances Visco *National Breast Cancer Coalition, Washington DC, USA.*
fmvisco@breastcancerdeadline2020.org

Cancer goal: vaccine is cause for optimism

As the originator of the project to develop a preventive breast-cancer vaccine, which is supported by the US National Breast Cancer Coalition's (NBCC) 2020 deadline, I question your dismissal of this new approach (*Nature* **491,** 637; 2012).

You base your pessimism on the genomic complexity of tumours and the length of time that clinical trials would take to test such a vaccine. But the genomic complexity apparent in mature tumours is not relevant for developing a preventive vaccine: the immune system needs only to be preactivated with antigens presented by the nascent tumour. Unstable RNA processing in a tumour is likely to be a richer source of these antigens than genomic mutations.

It could indeed take a decade or more to validate the treatment using today's clinical standards. But trials could be made shorter and less expensive if, for example, there were definitive ways to detect very early tumours.

You say that "discovery does not answer to deadlines", but accomplishment can. Given the alternatives, we should embrace this effort.

Stephen Albert Johnston

Arizona State University, Tempe, USA.
stephen.johnston@asu.edu

Tie carbon emissions to consumers

Global carbon dioxide emissions soared from 22.7 billion tonnes in 1990 to 33.9 billion tonnes last year, despite 20 years of attempted mitigation (*Nature* **491,** 656–658; 2012). The sizeable economic gaps between nations are largely responsible for the international deadlock in climate negotiations. A radically new approach is needed.

One solution would be to allocate common but differentiated responsibility for mitigating emissions to individual consumption activities, rather than to countries. Profligate consumers from both developed and emerging countries are the worst offenders for generating non-essential emissions. They should be held accountable for those emissions, irrespective of where they are generated.

Emissions produced after reallocating responsibility in this way could be quantified for nations, lifestyles or even every product consumed. This could be achieved by integrating the 'top-down' consumption-based accounting methods that are used to determine national total consumption emissions with 'bottom-up' carbon-footprint calculations based on analysis of products' life cycles (G. P. Peters *Curr. Opin. Environ. Sustain.* **2,** 245–250; 2010). A living standard, together with a per capita emission quota, would be defined so that people worldwide could meet their basic living requirements without mitigation costs.

Such standardized measures would allow new cap-and-trade policies and carbon-taxing mechanisms to run smoothly and effectively across different consumption groups at a global scale (see M. Grubb *Nature* **491,** 666–667; 2012).

Zhu Liu, Fengming Xi *Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang, China.*
liuzhu@iae.ac.cn
Dabo Guan *University of Leeds, Leeds, UK.*

Sexism: measure journal objectivity

Your initiative to address the issue of sexism in publishing is laudable (*Nature* **491,** 495; 2012). But scientific editors may be biased in other ways that influence the publication process — and they are not in the best position to recognize and correct their own biases.

An independent, external assessment body that regularly evaluates editorial practices across scientific journals might be the answer. This would promote transparency and reassure authors that their work is being dealt with fairly. An 'objectivity factor' resulting from such an assessment could become a key metric of journal performance, alongside its impact factor.

Boyan K. Garvalov *University of Giessen, Germany.*
boyan.garvalov@patho.med.uni-giessen.de

Sexism: control experimental bias

Implicit bias against women (*Nature* **491,** 495; 2012) has a record of influencing the design of experiments and collection of data in the life and mind sciences.

Over the past four decades, feminist scientists, historians and philosophers of science have presented case study after case study showing how sexist bias can distort scientific results. Examples include work by Rebecca Jordan-Young on brain-organization theory, Anne Fausto-Sterling on the biological study of sex differences and Anelis Kaiser on neuroimaging.

In the interest of greater accuracy, controls for sexist bias need to be more rigorous.

Letitia Meynell *Dalhousie University, Halifax,*
Nova Scotia, Canada.
letitia.meynell@dal.ca

Sexism: conferences should seek a balance

Conference committees and symposia chairs should follow *Nature*'s lead and be more open and proactive about gender balance (*Nature* **491,** 495; 2012).

Publishing an online declaration of gender-balance policy would help. Data from each of the past, say, three meetings could also be posted online to indicate the percentage of women who were registrants, invited plenary and keynote speakers, session speakers, programme committee members, or session chairs. If there is a significant gap between the percentage of women attending the conference and the overall percentage of women speakers or committee members, then a policy overhaul may be called for.

Jennifer L. Martin *University of Queensland, Brisbane, Australia.*
j.martin@imb.uq.edu.au

More trials needed to assess sleeping pills

Sam Fleishman suggests that the controlled use of sleeping pills helps to counter the life-disrupting consequences of insomnia (*Nature* **491,** 527; 2012). But, to our knowledge, there is no convincing evidence that sleep medication can reactivate the health-restoring functions of sleep.

Good-quality sleep improves cognitive performance, vigilance, memory and mood. Poor sleep is associated with potentially damaging physiological effects such as inflammation (M. R. Irwin *et al. Brain Behav. Immun.* **24,** 54–57; 2010), compromised immune-cell activity (E. Fondell *et al. Brain Behav. Immun.* **25,** 1367–1375; 2011) and telomere shortening (A. A. Prather *et al. J. Aging Res.* **2011,** 721390; 2011).

However, the regular use of sleeping pills has been linked with increased mortality and morbidity from infection, depression or cancer (see, for example, D. F. Kripke *et al. BMJ Open* **2,** e000850; 2012).

Large randomizing trials are needed to evaluate the benefits and risks of sleeping pills so that they can be prescribed more responsibly.

Maximilian Moser *Medical University of Graz, Austria*
max.moser@medunigraz.at
Daniel F. Kripke *University of California San Diego, California, USA.*

Toe–clipping vital to amphibian research

Keeping a record of the global decline of amphibian populations depends on the identification and marking of individuals in the field, commonly by toe-clipping. This work is under threat: the Brazilian federal government and non-governmental organizations want to prohibit toe-clipping without scientific justification.

Toe-clipping is a simple, cost-effective, minimally invasive marking technique that has been in use for decades in herpetological research. The Brazilian federal agency that enforces environmental policies, IBAMA, is claiming that the practice is a form of mutilation and should be a criminal offence under federal law (see go.nature.com/qkij7l; in Portuguese). The Brazilian Herpetological Society has protested against this anthropocentric position (see go.nature.com/thpwdg; in Portuguese) on the grounds that it would set back efforts to understand amphibian decline.

Brazil is home to the greatest amphibian diversity on Earth, but knowledge of population dynamics is scant and will remain so without access to reliable marking methods. We acknowledge that toe-clipping is not an ideal solution but, when properly implemented, it has minimal effects on amphibian survival and behaviour. Without this technique, it will be harder to obtain crucial information that could prevent amphibian species from becoming extinct (W. C. Funk *et al. Nature* **433,** 193; 2005).

Décio T. Corrêa* *State University of Campinas, São Paulo, Brazil.*
decio.tadeu@gmail.com
On behalf of 5 co-signatories (see go.nature.com/5gsgxo for full list).

Missing data mean holes in tree of life

As part of the Open Tree of Life project (http://opentreeoflife.org), we surveyed publications covering all domains of life and found that most phylogenetic trees and nucleotide alignments from the past two decades have been irrevocably lost.

Of 6,193 papers we surveyed in more than 100 peer-reviewed journals, only 17% present accessible trees and alignments (used to infer relatedness). Contacting lead authors to procure data sets was only 19% successful. DNA sequences were deposited in GenBank for almost all these studies, but it is the actual character alignments that are pivotal for reproducing phylogenetic analyses. We estimate that more than 64% of existing alignments or trees are permanently lost.

This problem will increasingly hinder phylogenetic inference as the use of whole-genome data sets becomes common. Journals need to reinforce a policy of online data deposition, either as supplementary material or in repositories such as TreeBASE (http://treebase.org) or Dryad (http://datadryad.org) — including for data sets based on previously published sequences. Ecologists, evolutionary biologists and others will then have access to rigorous phylogenetics for testing their hypotheses.

Bryan T. Drew* *University of Florida, Gainesville, USA.*
bdrew@ufl.edu
On behalf of 8 co-signatories (see go.nature.com/gzqcr9 for full list).

CORRECTION

The Outlook article 'Genetics: Searching for answers' (*Nature* **491** (suppl. 7422), S4–S6; 2012) incorrectly stated that the Simons Simplex Collection consists of 3,000 blood samples taken from more than 700 people. In fact, it contains 13,000 blood samples taken from 3,000 people.

Zhu Liu, Fengming Xi *Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang, China.*
liuzhu@iae.ac.cn
Dabo Guan *University of Leeds, Leeds, UK.*

Sexism: measure journal objectivity

Your initiative to address the issue of sexism in publishing is laudable (*Nature* **491,** 495; 2012). But scientific editors may be biased in other ways that influence the publication process — and they are not in the best position to recognize and correct their own biases.

An independent, external assessment body that regularly evaluates editorial practices across scientific journals might be the answer. This would promote transparency and reassure authors that their work is being dealt with fairly. An 'objectivity factor' resulting from such an assessment could become a key metric of journal performance, alongside its impact factor.

Boyan K. Garvalov *University of Giessen, Germany.*
boyan.garvalov@patho.med.uni-giessen.de

Sexism: control experimental bias

Implicit bias against women (*Nature* **491,** 495; 2012) has a record of influencing the design of experiments and collection of data in the life and mind sciences.

Over the past four decades, feminist scientists, historians and philosophers of science have presented case study after case study showing how sexist bias can distort scientific results. Examples include work by Rebecca Jordan-Young on brain-organization theory, Anne Fausto-Sterling on the biological study of sex differences and Anelis Kaiser on neuroimaging.

In the interest of greater accuracy, controls for sexist bias need to be more rigorous.

Letitia Meynell *Dalhousie University, Halifax,*
Nova Scotia, Canada.
letitia.meynell@dal.ca

Sexism: conferences should seek a balance

Conference committees and symposia chairs should follow *Nature*'s lead and be more open and proactive about gender balance (*Nature* **491,** 495; 2012).

Publishing an online declaration of gender-balance policy would help. Data from each of the past, say, three meetings could also be posted online to indicate the percentage of women who were registrants, invited plenary and keynote speakers, session speakers, programme committee members, or session chairs. If there is a significant gap between the percentage of women attending the conference and the overall percentage of women speakers or committee members, then a policy overhaul may be called for.

Jennifer L. Martin *University of Queensland, Brisbane, Australia.*
j.martin@imb.uq.edu.au

More trials needed to assess sleeping pills

Sam Fleishman suggests that the controlled use of sleeping pills helps to counter the life-disrupting consequences of insomnia (*Nature* **491,** 527; 2012). But, to our knowledge, there is no convincing evidence that sleep medication can reactivate the health-restoring functions of sleep.

Good-quality sleep improves cognitive performance, vigilance, memory and mood. Poor sleep is associated with potentially damaging physiological effects such as inflammation (M. R. Irwin *et al. Brain Behav. Immun.* **24,** 54–57; 2010), compromised immune-cell activity (E. Fondell *et al. Brain Behav. Immun.* **25,** 1367–1375; 2011) and telomere shortening (A. A. Prather *et al. J. Aging Res.* **2011,** 721390; 2011).

However, the regular use of sleeping pills has been linked with increased mortality and morbidity from infection, depression or cancer (see, for example, D. F. Kripke *et al. BMJ Open* **2,** e000850; 2012).

Large randomizing trials are needed to evaluate the benefits and risks of sleeping pills so that they can be prescribed more responsibly.

Maximilian Moser *Medical University of Graz, Austria*
max.moser@medunigraz.at
Daniel F. Kripke *University of California San Diego, California, USA.*

Toe–clipping vital to amphibian research

Keeping a record of the global decline of amphibian populations depends on the identification and marking of individuals in the field, commonly by toe-clipping. This work is under threat: the Brazilian federal government and non-governmental organizations want to prohibit toe-clipping without scientific justification.

Toe-clipping is a simple, cost-effective, minimally invasive marking technique that has been in use for decades in herpetological research. The Brazilian federal agency that enforces environmental policies, IBAMA, is claiming that the practice is a form of mutilation and should be a criminal offence under federal law (see go.nature.com/qkij7l; in Portuguese). The Brazilian Herpetological Society has protested against this anthropocentric position (see go.nature.com/thpwdg; in Portuguese) on the grounds that it would set back efforts to understand amphibian decline.

Brazil is home to the greatest amphibian diversity on Earth, but knowledge of population dynamics is scant and will remain so without access to reliable marking methods. We acknowledge that toe-clipping is not an ideal solution but, when properly implemented, it has minimal effects on amphibian survival and behaviour. Without this technique, it will be harder to obtain crucial information that could prevent amphibian species from becoming extinct (W. C. Funk *et al. Nature* **433,** 193; 2005).

Décio T. Corrêa* *State University of Campinas, São Paulo, Brazil.*
decio.tadeu@gmail.com
On behalf of 5 co-signatories (see go.nature.com/5gsqo for full list).

Missing data mean holes in tree of life

As part of the Open Tree of Life project (http://opentreeoflife.org), we surveyed publications covering all domains of life and found that most phylogenetic trees and nucleotide alignments from the past two decades have been irrevocably lost.

Of 6,193 papers we surveyed in more than 100 peer-reviewed journals, only 17% present accessible trees and alignments (used to infer relatedness). Contacting lead authors to procure data sets was only 19% successful. DNA sequences were deposited in GenBank for almost all these studies, but it is the actual character alignments that are pivotal for reproducing phylogenetic analyses. We estimate that more than 64% of existing alignments or trees are permanently lost.

This problem will increasingly hinder phylogenetic inference as the use of whole-genome data sets becomes common. Journals need to reinforce a policy of online data deposition, either as supplementary material or in repositories such as TreeBASE (http://treebase.org) or Dryad (http://datadryad.org) — including for data sets based on previously published sequences. Ecologists, evolutionary biologists and others will then have access to rigorous phylogenetics for testing their hypotheses.

Bryan T. Drew* *University of Florida, Gainesville, USA.*
bdrew@ufl.edu
On behalf of 8 co-signatories (see go.nature.com/gzqcr9 for full list).

CORRECTION

The Outlook article 'Genetics: Searching for answers' (*Nature* **491** (suppl. 7422), S4–S6; 2012) incorrectly stated that the Simons Simplex Collection consists of 3,000 blood samples taken from more than 700 people. In fact, it contains 13,000 blood samples taken from 3,000 people.

Zhu Liu, Fengming Xi *Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang, China.*
liuzhu@iae.ac.cn
Dabo Guan *University of Leeds, Leeds, UK.*

Sexism: measure journal objectivity

Your initiative to address the issue of sexism in publishing is laudable (*Nature* **491,** 495; 2012). But scientific editors may be biased in other ways that influence the publication process — and they are not in the best position to recognize and correct their own biases.

An independent, external assessment body that regularly evaluates editorial practices across scientific journals might be the answer. This would promote transparency and reassure authors that their work is being dealt with fairly. An 'objectivity factor' resulting from such an assessment could become a key metric of journal performance, alongside its impact factor.

Boyan K. Garvalov *University of Giessen, Germany.*
boyan.garvalov@patho.med.uni-giessen.de

Sexism: control experimental bias

Implicit bias against women (*Nature* **491,** 495; 2012) has a record of influencing the design of experiments and collection of data in the life and mind sciences.

Over the past four decades, feminist scientists, historians and philosophers of science have presented case study after case study showing how sexist bias can distort scientific results. Examples include work by Rebecca Jordan-Young on brain-organization theory, Anne Fausto-Sterling on the biological study of sex differences and Anelis Kaiser on neuroimaging.

In the interest of greater accuracy, controls for sexist bias need to be more rigorous.

Letitia Meynell *Dalhousie University, Halifax,*
Nova Scotia, Canada.
letitia.meynell@dal.ca

Sexism: conferences should seek a balance

Conference committees and symposia chairs should follow *Nature*'s lead and be more open and proactive about gender balance (*Nature* **491,** 495; 2012).

Publishing an online declaration of gender-balance policy would help. Data from each of the past, say, three meetings could also be posted online to indicate the percentage of women who were registrants, invited plenary and keynote speakers, session speakers, programme committee members, or session chairs. If there is a significant gap between the percentage of women attending the conference and the overall percentage of women speakers or committee members, then a policy overhaul may be called for.

Jennifer L. Martin *University of Queensland, Brisbane, Australia.*
j.martin@imb.uq.edu.au

More trials needed to assess sleeping pills

Sam Fleishman suggests that the controlled use of sleeping pills helps to counter the life-disrupting consequences of insomnia (*Nature* **491,** 527; 2012). But, to our knowledge, there is no convincing evidence that sleep medication can reactivate the health-restoring functions of sleep.

Good-quality sleep improves cognitive performance, vigilance, memory and mood. Poor sleep is associated with potentially damaging physiological effects such as inflammation (M. R. Irwin *et al. Brain Behav. Immun.* **24,** 54–57; 2010), compromised immune-cell activity (E. Fondell *et al. Brain Behav. Immun.* **25,** 1367–1375; 2011) and telomere shortening (A. A. Prather *et al. J. Aging Res.* **2011,** 721390; 2011).

However, the regular use of sleeping pills has been linked with increased mortality and morbidity from infection, depression or cancer (see, for example, D. F. Kripke *et al. BMJ Open* **2,** e000850; 2012).

Large randomizing trials are needed to evaluate the benefits and risks of sleeping pills so that they can be prescribed more responsibly.

Maximilian Moser *Medical University of Graz, Austria*
max.moser@medunigraz.at
Daniel F. Kripke *University of California San Diego, California, USA.*

Toe–clipping vital to amphibian research

Keeping a record of the global decline of amphibian populations depends on the identification and marking of individuals in the field, commonly by toe-clipping. This work is under threat: the Brazilian federal government and non-governmental organizations want to prohibit toe-clipping without scientific justification.

Toe-clipping is a simple, cost-effective, minimally invasive marking technique that has been in use for decades in herpetological research. The Brazilian federal agency that enforces environmental policies, IBAMA, is claiming that the practice is a form of mutilation and should be a criminal offence under federal law (see go.nature.com/qkij7l; in Portuguese). The Brazilian Herpetological Society has protested against this anthropocentric position (see go.nature.com/thpwdg; in Portuguese) on the grounds that it would set back efforts to understand amphibian decline.

Brazil is home to the greatest amphibian diversity on Earth, but knowledge of population dynamics is scant and will remain so without access to reliable marking methods. We acknowledge that toe-clipping is not an ideal solution but, when properly implemented, it has minimal effects on amphibian survival and behaviour. Without this technique, it will be harder to obtain crucial information that could prevent amphibian species from becoming extinct (W. C. Funk *et al. Nature* **433,** 193; 2005).

Décio T. Corrêa* *State University of Campinas, São Paulo, Brazil.*
decio.tadeu@gmail.com
On behalf of 5 co-signatories (see go.nature.com/5gsqo for full list).

Missing data mean holes in tree of life

As part of the Open Tree of Life project (http://opentreeoflife.org), we surveyed publications covering all domains of life and found that most phylogenetic trees and nucleotide alignments from the past two decades have been irrevocably lost.

Of 6,193 papers we surveyed in more than 100 peer-reviewed journals, only 17% present accessible trees and alignments (used to infer relatedness). Contacting lead authors to procure data sets was only 19% successful. DNA sequences were deposited in GenBank for almost all these studies, but it is the actual character alignments that are pivotal for reproducing phylogenetic analyses. We estimate that more than 64% of existing alignments or trees are permanently lost.

This problem will increasingly hinder phylogenetic inference as the use of whole-genome data sets becomes common. Journals need to reinforce a policy of online data deposition, either as supplementary material or in repositories such as TreeBASE (http://treebase.org) or Dryad (http://datadryad.org) — including for data sets based on previously published sequences. Ecologists, evolutionary biologists and others will then have access to rigorous phylogenetics for testing their hypotheses.

Bryan T. Drew* *University of Florida, Gainesville, USA.*
bdrew@ufl.edu
On behalf of 8 co-signatories (see go.nature.com/gzqcr9 for full list).

CORRECTION

The Outlook article 'Genetics: Searching for answers' (*Nature* **491** (suppl. 7422), S4–S6; 2012) incorrectly stated that the Simons Simplex Collection consists of 3,000 blood samples taken from more than 700 people. In fact, it contains 13,000 blood samples taken from 3,000 people.

Zhu Liu, Fengming Xi *Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang, China.*
liuzhu@iae.ac.cn
Dabo Guan *University of Leeds, Leeds, UK.*

Sexism: measure journal objectivity

Your initiative to address the issue of sexism in publishing is laudable (*Nature* **491,** 495; 2012). But scientific editors may be biased in other ways that influence the publication process — and they are not in the best position to recognize and correct their own biases.

An independent, external assessment body that regularly evaluates editorial practices across scientific journals might be the answer. This would promote transparency and reassure authors that their work is being dealt with fairly. An 'objectivity factor' resulting from such an assessment could become a key metric of journal performance, alongside its impact factor.

Boyan K. Garvalov *University of Giessen, Germany.*
boyan.garvalov@patho.med.uni-giessen.de

Sexism: control experimental bias

Implicit bias against women (*Nature* **491,** 495; 2012) has a record of influencing the design of experiments and collection of data in the life and mind sciences.

Over the past four decades, feminist scientists, historians and philosophers of science have presented case study after case study showing how sexist bias can distort scientific results. Examples include work by Rebecca Jordan-Young on brain-organization theory, Anne Fausto-Sterling on the biological study of sex differences and Anelis Kaiser on neuroimaging.

In the interest of greater accuracy, controls for sexist bias need to be more rigorous.

Letitia Meynell *Dalhousie University, Halifax,*

Nova Scotia, Canada.
letitia.meynell@dal.ca

Sexism: conferences should seek a balance

Conference committees and symposia chairs should follow *Nature*'s lead and be more open and proactive about gender balance (*Nature* **491,** 495; 2012).

Publishing an online declaration of gender-balance policy would help. Data from each of the past, say, three meetings could also be posted online to indicate the percentage of women who were registrants, invited plenary and keynote speakers, session speakers, programme committee members, or session chairs. If there is a significant gap between the percentage of women attending the conference and the overall percentage of women speakers or committee members, then a policy overhaul may be called for.

Jennifer L. Martin *University of Queensland, Brisbane, Australia.*
j.martin@imb.uq.edu.au

More trials needed to assess sleeping pills

Sam Fleishman suggests that the controlled use of sleeping pills helps to counter the life-disrupting consequences of insomnia (*Nature* **491,** 527; 2012). But, to our knowledge, there is no convincing evidence that sleep medication can reactivate the health-restoring functions of sleep.

Good-quality sleep improves cognitive performance, vigilance, memory and mood. Poor sleep is associated with potentially damaging physiological effects such as inflammation (M. R. Irwin *et al. Brain Behav. Immun.* **24,** 54–57; 2010), compromised immune-cell activity (E. Fondell *et al. Brain Behav. Immun.* **25,** 1367–1375; 2011) and telomere shortening (A. A. Prather *et al. J. Aging Res.* **2011,** 721390; 2011).

However, the regular use of sleeping pills has been linked with increased mortality and morbidity from infection, depression or cancer (see, for example, D. F. Kripke *et al. BMJ*

Open **2,** e000850; 2012).

Large randomizing trials are needed to evaluate the benefits and risks of sleeping pills so that they can be prescribed more responsibly.

Maximilian Moser *Medical University of Graz, Austria*
max.moser@medunigraz.at
Daniel F. Kripke *University of California San Diego, California, USA.*

Toe–clipping vital to amphibian research

Keeping a record of the global decline of amphibian populations depends on the identification and marking of individuals in the field, commonly by toe-clipping. This work is under threat: the Brazilian federal government and non-governmental organizations want to prohibit toe-clipping without scientific justification.

Toe-clipping is a simple, cost-effective, minimally invasive marking technique that has been in use for decades in herpetological research. The Brazilian federal agency that enforces environmental policies, IBAMA, is claiming that the practice is a form of mutilation and should be a criminal offence under federal law (see go.nature.com/qkij7l; in Portuguese). The Brazilian Herpetological Society has protested against this anthropocentric position (see go.nature.com/thpwdg; in Portuguese) on the grounds that it would set back efforts to understand amphibian decline.

Brazil is home to the greatest amphibian diversity on Earth, but knowledge of population dynamics is scant and will remain so without access to reliable marking methods. We acknowledge that toe-clipping is not an ideal solution but, when properly implemented, it has minimal effects on amphibian survival and behaviour. Without this technique, it will be harder to obtain crucial information that could prevent amphibian species from becoming extinct (W. C. Funk *et al. Nature* **433,** 193; 2005).

Décio T. Corrêa* *State University of Campinas, São Paulo, Brazil.*
decio.tadeu@gmail.com

*On behalf of 5 co-signatories (see go.nature.com/5sgsqo for full list).

Missing data mean holes in tree of life

As part of the Open Tree of Life project (http://opentreeoflife.org), we surveyed publications covering all domains of life and found that most phylogenetic trees and nucleotide alignments from the past two decades have been irrevocably lost.

Of 6,193 papers we surveyed in more than 100 peer-reviewed journals, only 17% present accessible trees and alignments (used to infer relatedness). Contacting lead authors to procure data sets was only 19% successful. DNA sequences were deposited in GenBank for almost all these studies, but it is the actual character alignments that are pivotal for reproducing phylogenetic analyses. We estimate that more than 64% of existing alignments or trees are permanently lost.

This problem will increasingly hinder phylogenetic inference as the use of whole-genome data sets becomes common. Journals need to reinforce a policy of online data deposition, either as supplementary material or in repositories such as TreeBASE (http://treebase.org) or Dryad (http://datadryad.org) — including for data sets based on previously published sequences. Ecologists, evolutionary biologists and others will then have access to rigorous phylogenetics for testing their hypotheses.

Bryan T. Drew* *University of Florida, Gainesville, USA.*
bdrew@ufl.edu
*On behalf of 8 co-signatories (see go.nature.com/gzqcr9 for full list).

CORRECTION

The Outlook article 'Genetics: Searching for answers' (*Nature* **491** (suppl. 7422), S4–S6; 2012) incorrectly stated that the Simons Simplex Collection consists of 3,000 blood samples taken from more than 700 people. In fact, it contains 13,000 blood samples taken from 3,000 people.

Zhu Liu, Fengming Xi *Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang, China.*
liuzhu@iae.ac.cn
Dabo Guan *University of Leeds, Leeds, UK.*

Sexism: measure journal objectivity

Your initiative to address the issue of sexism in publishing is laudable (*Nature* **491,** 495; 2012). But scientific editors may be biased in other ways that influence the publication process — and they are not in the best position to recognize and correct their own biases.

An independent, external assessment body that regularly evaluates editorial practices across scientific journals might be the answer. This would promote transparency and reassure authors that their work is being dealt with fairly. An 'objectivity factor' resulting from such an assessment could become a key metric of journal performance, alongside its impact factor.

Boyan K. Garvalov *University of Giessen, Germany.*
boyan.garvalov@patho.med.uni-giessen.de

Sexism: control experimental bias

Implicit bias against women (*Nature* **491,** 495; 2012) has a record of influencing the design of experiments and collection of data in the life and mind sciences.

Over the past four decades, feminist scientists, historians and philosophers of science have presented case study after case study showing how sexist bias can distort scientific results. Examples include work by Rebecca Jordan-Young on brain-organization theory, Anne Fausto-Sterling on the biological study of sex differences and Anelis Kaiser on neuroimaging.

In the interest of greater accuracy, controls for sexist bias need to be more rigorous.

Letitia Meynell *Dalhousie University, Halifax,*

Nova Scotia, Canada.
letitia.meynell@dal.ca

Sexism: conferences should seek a balance

Conference committees and symposia chairs should follow *Nature*'s lead and be more open and proactive about gender balance (*Nature* **491,** 495; 2012).

Publishing an online declaration of gender-balance policy would help. Data from each of the past, say, three meetings could also be posted online to indicate the percentage of women who were registrants, invited plenary and keynote speakers, session speakers, programme committee members, or session chairs. If there is a significant gap between the percentage of women attending the conference and the overall percentage of women speakers or committee members, then a policy overhaul may be called for.

Jennifer L. Martin *University of Queensland, Brisbane, Australia.*
j.martin@imb.uq.edu.au

More trials needed to assess sleeping pills

Sam Fleishman suggests that the controlled use of sleeping pills helps to counter the life-disrupting consequences of insomnia (*Nature* **491,** 527; 2012). But, to our knowledge, there is no convincing evidence that sleep medication can reactivate the health-restoring functions of sleep.

Good-quality sleep improves cognitive performance, vigilance, memory and mood. Poor sleep is associated with potentially damaging physiological effects such as inflammation (M. R. Irwin *et al. Brain Behav. Immun.* **24,** 54–57; 2010), compromised immune-cell activity (E. Fondell *et al. Brain Behav. Immun.* **25,** 1367–1375; 2011) and telomere shortening (A. A. Prather *et al. J. Aging Res.* **2011,** 721390; 2011).

However, the regular use of sleeping pills has been linked with increased mortality and morbidity from infection, depression or cancer (see, for example, D. F. Kripke *et al. BMJ*

Open **2,** e000850; 2012).

Large randomizing trials are needed to evaluate the benefits and risks of sleeping pills so that they can be prescribed more responsibly.

Maximilian Moser *Medical University of Graz, Austria*
max.moser@medunigraz.at
Daniel F. Kripke *University of California San Diego, California, USA.*

Toe–clipping vital to amphibian research

Keeping a record of the global decline of amphibian populations depends on the identification and marking of individuals in the field, commonly by toe-clipping. This work is under threat: the Brazilian federal government and non-governmental organizations want to prohibit toe-clipping without scientific justification.

Toe-clipping is a simple, cost-effective, minimally invasive marking technique that has been in use for decades in herpetological research. The Brazilian federal agency that enforces environmental policies, IBAMA, is claiming that the practice is a form of mutilation and should be a criminal offence under federal law (see go.nature.com/qkij7l; in Portuguese). The Brazilian Herpetological Society has protested against this anthropocentric position (see go.nature.com/thpwdg; in Portuguese) on the grounds that it would set back efforts to understand amphibian decline.

Brazil is home to the greatest amphibian diversity on Earth, but knowledge of population dynamics is scant and will remain so without access to reliable marking methods. We acknowledge that toe-clipping is not an ideal solution but, when properly implemented, it has minimal effects on amphibian survival and behaviour. Without this technique, it will be harder to obtain crucial information that could prevent amphibian species from becoming extinct (W. C. Funk *et al. Nature* **433,** 193; 2005).

Décio T. Corrêa* *State University of Campinas, São Paulo, Brazil.*
decio.tadeu@gmail.com

On behalf of 5 co-signatories (see go.nature.com/5gsqo for full list).

Missing data mean holes in tree of life

As part of the Open Tree of Life project (http://opentreeoflife.org), we surveyed publications covering all domains of life and found that most phylogenetic trees and nucleotide alignments from the past two decades have been irrevocably lost.

Of 6,193 papers we surveyed in more than 100 peer-reviewed journals, only 17% present accessible trees and alignments (used to infer relatedness). Contacting lead authors to procure data sets was only 19% successful. DNA sequences were deposited in GenBank for almost all these studies, but it is the actual character alignments that are pivotal for reproducing phylogenetic analyses. We estimate that more than 64% of existing alignments or trees are permanently lost.

This problem will increasingly hinder phylogenetic inference as the use of whole-genome data sets becomes common. Journals need to reinforce a policy of online data deposition, either as supplementary material or in repositories such as TreeBASE (http://treebase.org) or Dryad (http://datadryad.org) — including for data sets based on previously published sequences. Ecologists, evolutionary biologists and others will then have access to rigorous phylogenetics for testing their hypotheses.

Bryan T. Drew* *University of Florida, Gainesville, USA.*
bdrew@ufl.edu
On behalf of 8 co-signatories (see go.nature.com/gzqcr9 for full list).

CORRECTION

The Outlook article 'Genetics: Searching for answers' (*Nature* **491** (suppl. 7422), S4–S6; 2012) incorrectly stated that the Simons Simplex Collection consists of 3,000 blood samples taken from more than 700 people. In fact, it contains 13,000 blood samples taken from 3,000 people.

ARTICLE

Autism–related deficits via dysregulated eIF4E–dependent translational control

Christos G. Gkogkas[1], Arkady Khoutorsky[1]*, Israeli Ran[2]*, Emmanouil Rampakakis[1,3], Tatiana Nevarko[1], Daniel B. Weatherill[2], Cristina Vasuta[2], Stephanie Yee[4], Morgan Truitt[5], Paul Dallaire[6], François Major[6], Paul Lasko[4], Davide Ruggero[5], Karim Nader[7], Jean-Claude Lacaille[2] & Nahum Sonenberg[1]

Hyperconnectivity of neuronal circuits due to increased synaptic protein synthesis is thought to cause autism spectrum disorders (ASDs). The mammalian target of rapamycin (mTOR) is strongly implicated in ASDs by means of upstream signalling; however, downstream regulatory mechanisms are ill–defined. Here we show that knockout of the eukaryotic translation initiation factor 4E-binding protein 2 (4E-BP2)—an eIF4E repressor downstream of mTOR—or eIF4E overexpression leads to increased translation of neuroligins, which are postsynaptic proteins that are causally linked to ASDs. Mice that have the gene encoding 4E-BP2 (*Eif4ebp2*) knocked out exhibit an increased ratio of excitatory to inhibitory synaptic inputs and autistic–like behaviours (that is, social interaction deficits, altered communication and repetitive/stereotyped behaviours). Pharmacological inhibition of eIF4E activity or normalization of neuroligin 1, but not neuroligin 2, protein levels restores the normal excitation/inhibition ratio and rectifies the social behaviour deficits. Thus, translational control by eIF4E regulates the synthesis of neuroligins, maintaining the excitation–to–inhibition balance, and its dysregulation engenders ASD–like phenotypes.

ASDs encompass unique behaviours within three domains: social interactions, communication and repetitive or restricted interests and behaviours[1]. Dysregulation of the molecular machinery controlling synaptic messenger RNA translation has been postulated to cause ASD, mainly through increased synthesis of synaptic proteins and consequently augmented connectivity[2].

A key factor required for translation initiation of the vast majority of mRNAs is the eIF4F complex that includes the cap-binding protein eIF4E, the RNA helicase eIF4A and the modular scaffolding protein eIF4G that bridges the mRNA to the ribosome[3]. 4E-BP proteins repress translation initiation by disrupting eIF4F complex formation[4]. Although most cellular mRNAs require basal amounts of eIF4E to be translated, eIF4E preferentially enhances translation of a selective group of mRNAs with extensive secondary structure at their 5′ untranslated regions (UTRs) or other sequence elements[3,5].

mTOR exists in two complexes: mTORC1, which is sensitive to the drug rapamycin, and mTORC2, which is not[3]. mTORC1 phosphorylates 4E-BPs[4], stimulating translation. Hypophosphorylated forms of 4E-BPs bind to eIF4E and inhibit translation initiation. There are three mammalian *4E-BP* paralogues encoding 4E-BP1, 4E-BP2 and 4E-BP3. 4E-BP2 is the major form in the mammalian brain and has an important role in long-lasting synaptic plasticity, learning and memory[6].

Upstream of mTOR, germline mutations in the phosphatase and tensin homologue on chromosome ten (*PTEN*) are present in about 1–5% of patients with autism[7]. *Pten* knockout mice exhibit cognitive impairment and deficits in social interactions[8], which are rescued by rapamycin[9]. Mutations in the two tuberous sclerosis complex (*TSC1* and *TSC2*) genes cause autism in a subset of patients[10]. *Tsc2*[+/−] mice display synaptic plasticity and memory deficits[11,12], which are rescued by rapamycin, and enhanced ultrasonic vocalizations (USVs) in mouse pups[13]. Deletion of *Tsc1* in cerebellar Purkinje cells in mice

leads to autism-like phenotypes, which are rescued by rapamycin[14]. Recently, a *de novo* mutated eIF4GI was reported as a component of a highly interconnected network of proteins linked with sporadic autism[15]. Moreover, the mTORC1–eIF4E pathway is hyperactivated in fragile-X syndrome (FXS) patients diagnosed with autism[16]. The hypothesis that eIF4E dysregulation is causally linked to ASD is consistent with recent findings that CYFIP1 (cytoplasmic fragile-X mental retardation protein (FMRP) interacting protein 1), an eIF4E binding protein[17], is associated with ASD[18].

Aberrant information processing owing to the increased ratio of synaptic excitation to inhibition (E/I) has been proposed to cause ASD[19,20]. Children with ASD exhibit elevations in resting-state neuronal activity, supporting an E/I imbalance as a neurobiological feature of ASD[21]. Modulations of the E/I balance using optogenetics directly influence autism-like behaviours in adult mice[22]. Several ASD mouse models display E/I imbalances due to altered glutamatergic excitation or to altered GABA (γ-aminobutyric acid)-ergic inhibition, leading to an increased[23,24] or decreased[25–28] E/I ratio. The balance of excitatory and inhibitory synapses is largely controlled by the expression of adhesion molecules (neuroligins (NLGNs), neurexins) and scaffolding proteins (postsynaptic density protein 95 (PSD95), gephyrin) in neurons[29]. Overexpression, knockout or knock-in of ASD-related neuroligin mutations reveals ASD-like phenotypes with alterations in the E/I balance (summarized in Supplementary Table 1).

A causal link between dysregulated eIF4E-dependent translation and the development of ASD is lacking. Here we show that deletion of mouse *Eif4ebp2* leads to autistic-like behaviours, including social interaction deficits, altered communication and repetitive/stereotyped behaviours. We find that translation of neuroligin mRNAs is enhanced in *Eif4ebp2* knockout and eIF4E-overexpressing mice. Furthermore, unitary excitatory and inhibitory synaptic activity is increased in

[1]Department of Biochemistry & Goodman Cancer Research Centre, McGill University, Montreal, Quebec H3A 1A3, Canada. [2]GRSNC and Department of Physiology, Université de Montréal, Montreal, Quebec H3C 3J7, Canada. [3]JSS Medical Research Inc., Montreal, Quebec H4S 1N8, Canada. [4]Department of Biology, McGill University, Montreal, Quebec H3G 0B1, Canada. [5]School of Medicine and Department of Urology, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California 94158, USA. [6]Institute for Research in Immunology and Cancer, and Department of Computer Science, Université de Montréal, Montreal, Quebec H3C 3J7, Canada. [7]Department of Psychology, McGill University, Montreal, Quebec H3A 1B1, Canada.
*These authors contributed equally to this work.

Eif4ebp2 knockout mice, with excitation predominantly being augmented. Moreover, pharmacological normalization of eIF4F activity corrects the E/I imbalance and rescues the social behaviour deficits of *Eif4ebp2* knockout mice. Knockdown of *Nlgn1*, but not *Nlgn2*, in *Eif4ebp2* knockout mice reverses the defects in social interactions and rectifies the E/I imbalance. Thus, enhanced eIF4E-dependent translation of mRNAs encoding synaptic adhesion proteins, such as neuroligins, leads to an aberrant increase in excitatory synaptic function and results in ASD-like phenotypes.

Autistic-like phenotypes in *Eif4ebp2* knockout mice

As *Eif4ebp2* deletion causes increased eIF4F complex formation and activity, we postulated that it could engender ASD-related behaviours. Social interaction deficits are a salient autistic behavioural feature in humans[1]. To test our hypothesis, we used a three-chamber social arena to assess the preference of a test mouse for a social (stranger1) over a non-social (empty wire cage) stimulus or for social novelty (stranger2). In comparison with wild-type littermates, *Eif4ebp2* knockout mice displayed impaired social approach behaviour as depicted by the significantly reduced time spent in the stranger1 compartment and the decreased interaction time with stranger1 (Fig. 1a and Supplementary Fig. 2). Given the lack of preference for the social stimulus in the first portion of the three-chamber test, preference for social novelty was also reduced in *Eif4ebp2* knockout mice (Supplementary Fig. 3). Because the reduced number of entries of the *Eif4ebp2* knockout mice into the stranger1 chamber relative to wild-type littermates (Fig. 1a) might underlie an anxiety-like phenotype, we assessed anxiety levels. No differences between wild-type and *Eif4ebp2* knockout mice were found using the elevated plus maze test following the three-chamber social approach test (Supplementary Fig. 4d). Moreover, wild-type and *Eif4ebp2* knockout mice were

indistinguishable in their initial exploration of the three-chamber arena (Supplementary Fig. 4a).

To characterize the social behaviour of the *Eif4ebp2* knockout mice further, we performed reciprocal social interaction tests: (1) in their home cage and (2) in a separate cage (Fig. 1b, c). Pairs of wild-type–wild-type, knockout–wild-type or knockout–knockout mice were either recorded in their home cage, or introduced to a familiar open-field environment. The knockout–wild-type and knockout–knockout pairs interacted for a shorter period of time, as compared to the wild-type–wild-type pair (Fig. 1b, c), without any difference in the total number of contacts between the wild-type–wild-type or wild-type–knockout pairs of mice. Knockout–knockout pairs initiated significantly fewer contacts (Fig. 1b, c). This discrepancy can be attributed to the larger number of contacts initiated by the wild-type mice in the wild-type–knockout pairs (Supplementary Fig. 2h). Anxiety is not a confounding factor for these behaviours, as demonstrated by the similar anxiety levels in wild-type and *Eif4ebp2* knockout mice in the elevated plus maze test (Supplementary Fig. 4b, c). To assess stereotyped/repetitive behaviours, we recorded self-grooming and used a marble-burying test. *Eif4ebp2* knockout mice displayed longer bouts and total duration of self-grooming (Fig. 1d and Supplementary Fig. 2i) and buried significantly more marbles than wild-type littermates (Fig. 1e).

Isolation-induced USVs emitted from mouse pups separated from their mother are a quantifiable behaviour considered relevant to ASD[30]. $Tsc2^{+/-}$ pups display enhanced USVs (increased number of calls, duration, peak amplitude and peak frequency)[13,14]. We recorded USVs in *Eif4ebp2* knockout and wild-type littermate pups after separation from their dams (Fig. 1f–h). *Eif4ebp2* knockout pups emitted a higher number of calls (Fig. 1f) with a longer duration (Fig. 1g) at all postnatal days relative to wild-type littermates. Furthermore, the
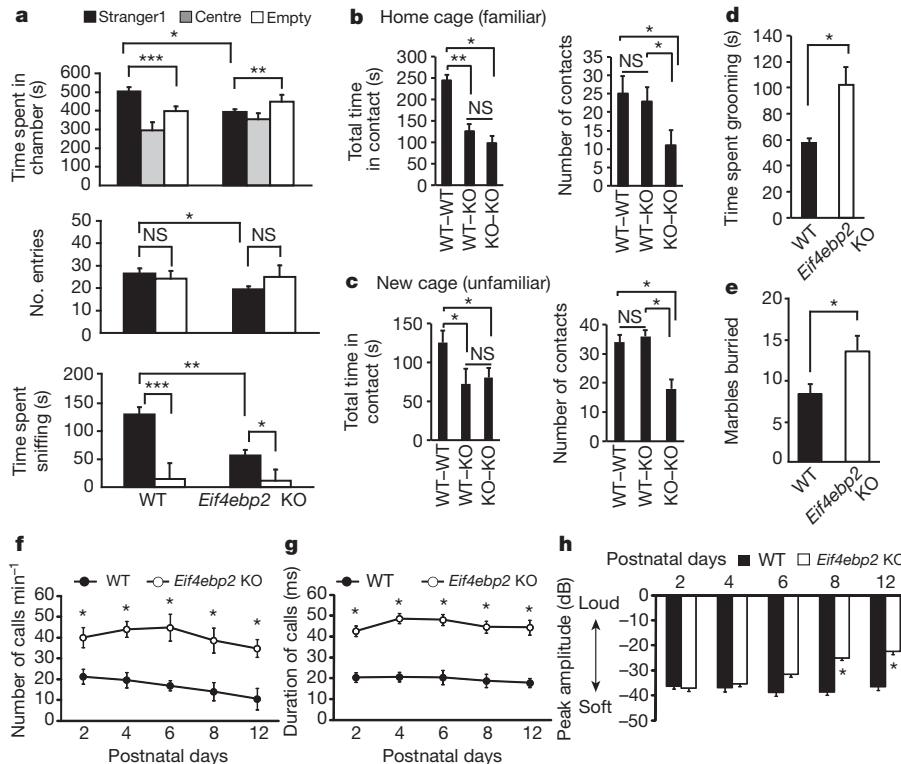
Figure 1 | Social interaction deficits, repetitive behaviour and elevated USVs in *Eif4ebp2* knockout mice. a, Three-chamber social interaction test showing time spent in each chamber, number of entries and time spent sniffing the wire cage; $n = 12$ for each group; ***$P < 0.001$, **$P < 0.01$, *$P < 0.03$, two-way analysis of variance (ANOVA) with Bonferroni's post-hoc test. KO, knockout; NS, not significant; WT, wild type. **b, c**, Home cage (**b**) and reciprocal social interactions (**c**) tests showing total time in contact and number of

contacts; $n = 18$ per group; **$P < 0.02$, *$P < 0.03$; one way ANOVA with Bonferroni's post-hoc test. **d**, Self-grooming test showing total time spent grooming. **e**, Marble-burying test showing marbles buried; $n = 12$ per group; *$P < 0.03$; Student's *t*-test. **f–h**, Isolation-induced USVs at various postnatal days. Shown are number of calls per min (**f**), duration of calls (**g**) and amplitude of USVs (**h**) in pups; $n = 12$ per group; *$P < 0.05$; two-way ANOVA, with Bonferroni's post-hoc test. All error bars show mean ± s.e.m.

peak amplitude of the USVs was higher on postnatal days 8 and 12 (Fig. 1h). Taken together, our data demonstrate that genetic deletion of *Eif4ebp2*, mimicking eIF4E upregulation, engenders ASD-like behaviours.

Increased eIF4E–dependent translation of neuroligins

We hypothesized that the ASD-like phenotypes of the *Eif4ebp2* knock-out mice are a result of altered translation of a subset of mRNAs, the initiation of which is stimulated by eIF4E activity. To test this hypothesis, we performed polysome profiling, which measures translation initiation rates of hippocampal lysates from *Eif4ebp2* knockout and eIF4E-overexpressing (βT-*Eif4e*)[31] mice. The polysome profile was not significantly altered in *Eif4ebp2* knockout mice as compared to wild-type littermates (Fig. 2a and Supplementary Fig. 5d). This is in agreement with the lack of changes in [35S]Met/Cys incorporation (Supplementary Fig. 5a, b) and previous reports[32]. We then examined the polysome distribution of 24 mRNAs coding for proteins known to be associated with ASD (Supplementary Table 2)[33,34]. These include mRNAs encoding the family of neuroligins and neurexins, SHANK2 and SHANK3 (SH3 and multiple ankyrin repeat domains protein) and MeCP2 (methyl-CpG-binding protein). Notably, out of the 24 mRNAs examined, only mRNA for *Nlgn1*, *Nlgn2*, *Nlgn3* and *Nlgn4* was shifted towards heavier polysome fractions in the *Eif4ebp2* knockout mice relative to wild-type littermates (Fig. 2c), indicating increased translation initiation. Neither mRNA levels of *Nlgn1–Nlgn4* nor those of any other mRNA were different between *Eif4ebp2* knockout and wild-type mice (Supplementary Fig. 6c), thus excluding transcriptional effects.

As expected, protein amounts of all four neuroligins were increased in crude and synaptosomal extracts from *Eif4ebp2* knockout hippocampi (Fig. 2e, g). Elevated translation, but not transcription, of neuroligin mRNAs was also observed in polysome profiles of βT-*Eif4e* mice (Fig. 2d and Supplementary Fig. 6c) with no differences in translation or transcription of other mRNAs (Supplementary Figs 6c and 7c). Accordingly, protein levels of all four neuroligins were increased in the synaptosomal fractions from βT-*Eif4e* mice relative to wild type (Fig. 2f, h). The increased translation of neuroligins, observed only in synaptosomal fractions of βT-*Eif4e* mice, can be explained by the mild overexpression (by 29%) of eIF4E (Fig. 2f). In contrast, no changes were observed in the translation of mRNAs encoding neurexins (Supplementary Fig. 6) or PSD95 (*Dlg4* gene; Fig. 2c–f) in *Eif4ebp2* knockout or βT-*Eif4e* versus wild-type mice. Translation of mRNAs coding for gephyrin (*Gphn*), SHANK2, SHANK3 and SAPAP3 (PSD95/synapse associated protein 90-binding protein 3) (*Dlgap3*) proteins, which orchestrate the formation of the postsynaptic scaffolding[29], was not changed in *Eif4ebp2* knockout or in the βT-*Eif4e* mice (Supplementary Fig. 7). In conclusion, relief of translational suppression by loss of 4E-BP2 or by overexpression of eIF4E selectively enhances the synthesis of neuroligins, causing an imbalance between these adhesion proteins relative to scaffolding proteins.

Increased synaptic activity in *Eif4ebp2* knockouts

An altered E/I ratio was reported in several ASD mouse models[23–28]. Therefore, we examined the effects of *Eif4ebp2* knockout on unitary excitatory and inhibitory synaptic transmission in acute hippocampal
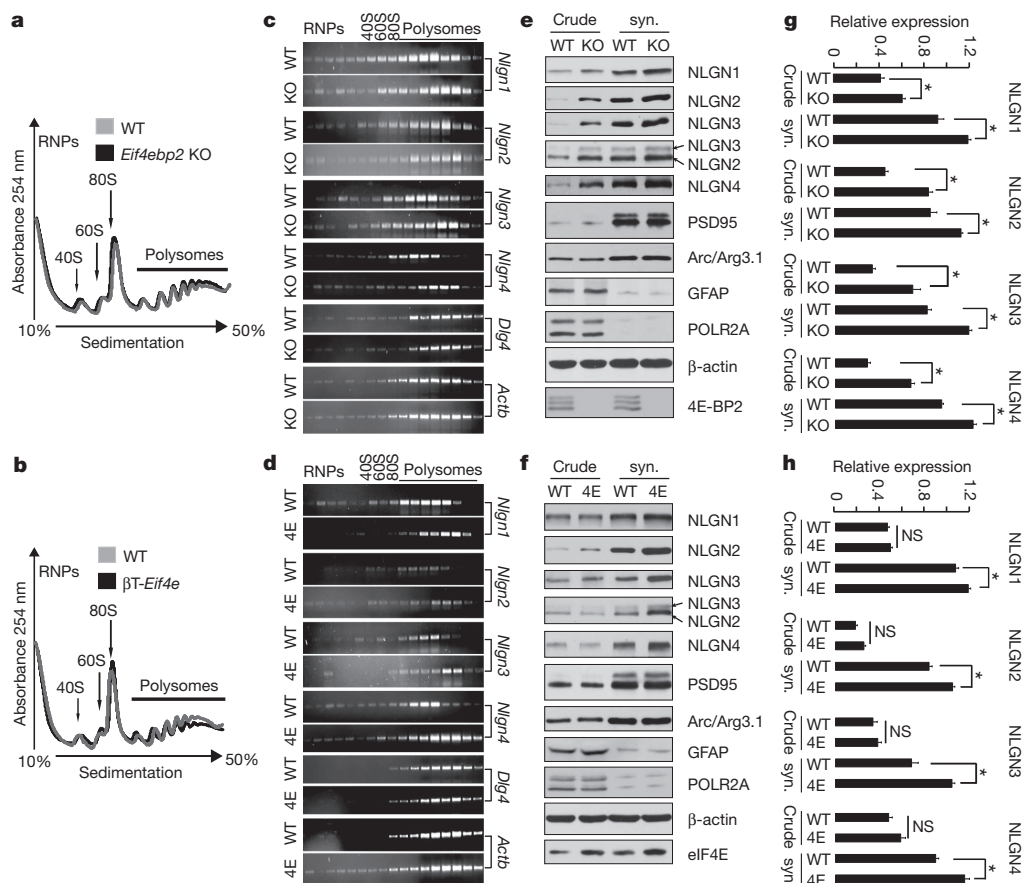
Figure 2 | Enhanced eIF4E-dependent translation of neuroligin mRNAs.
a, **b**, Polysome profiles from hippocampal lysates of wild-type and *Eif4ebp2* knockout mice (**a**) and wild-type and βT-*Eif4e* mice (**b**). Positions of 40S, 60S and 80S ribosome peaks and polysomes are indicated. **c**, **d**, RT–PCR on RNA extracted from polysome fractions from wild-type and *Eif4ebp2* knockout (**c**) and wild-type and βT-*Eif4e* (**d**; 4E) hippocampi. Representative gel images

are shown; *n* = 4 per group. **e**, **f**, Representative immunoblots of crude and synaptosomal (syn.) fractions from hippocampal lysates of wild-type or *Eif4ebp2* knockout (**e**) and wild-type or βT-*Eif4e* (**f**; 4E) mice, probed with antibodies against the indicated proteins. **g**, **h**, Quantification of immunoblots from **e** and **f** for neuroligins; *n* = 4 per group; *P < 0.05; Student's *t*-test. All error bars show mean ± s.e.m.

slices. AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) receptor-mediated miniature excitatory postsynaptic currents (mEPSCs) were increased both in amplitude and frequency in pyramidal neurons of *Eif4ebp2* knockout relative to wild-type mice (Fig. 3a, b), indicating that 4E-BP2 de-repression augments transmission at excitatory synapses in acute hippocampal slices. We next examined whether transmission was similarly altered at inhibitory synapses. GABA$_A$ receptor-mediated miniature inhibitory postsynaptic currents (mIPSCs) were significantly greater in amplitude, but not in frequency (unlike mEPSCs), in pyramidal cells of *Eif4ebp2* knockout relative to wild-type slices (Fig. 3c, d). To study the differential effect on excitatory versus inhibitory transmission, we measured the total synaptic charge transfer for mEPSCs and mIPSCs over a 10-min period—a parameter reflecting both amplitude and frequency of miniature synaptic events. The total charge transfer was significantly increased for both mEPSCs and mIPSCs in *Eif4ebp2* knockout pyramidal cells (Fig. 3e). However, the normalized increase of total charge transfer in *Eif4ebp2* knockout relative to wild-type mice was significantly larger for mEPSCs than for mIPSCs (Fig. 3f). We also recorded minimal stimulation-evoked EPSCs and IPSCs from CA1 pyramidal neurons in wild-type and *Eif4ebp2* knockout slices and observed an increase in both EPSC and IPSC amplitude, but no change in paired-pulse ratio (Supplementary Fig. 8). Moreover, we observed a greater increase in vesicular glutamate transporter (VGLUT) than vesicular GABA transporter (VGAT) puncta and protein amounts in the CA1 dendritic layer of *Eif4ebp2* knockout hippocampal slices, relative to wild-type littermates (Supplementary Fig. 9a–g). Finally, spine density is increased in

CA1 pyramidal cell dendrites in *Eif4ebp2* knockout mice (Supplementary Fig. 9h, i). Thus, 4E-BP2 de-repression results in altered E/I balance.

Pharmacological rescue of ASD–like phenotypes

To bolster the conclusion that translation initiation through eIF4E is critical in altering synaptic function and thus causing ASD, we used the selective inhibitor 4EGI-1 (ref. 35), which prevents eIF4E binding to eIF4G, to test whether it would rescue autistic-like behaviours in *Eif4ebp2* knockout mice. First, we performed cap-column pull-down assays with hippocampal lysates to assess the status of the eIF4F complex. As expected, increased amounts of eIF4G and eIF4E were bound to the m^{7}GDP-cap analogue matrix in *Eif4ebp2* knockout mice, as compared to wild-type mice (Fig. 4b). Infusion of 4EGI-1 reduced the amounts of cap-bound eIF4G and eIF4E to wild-type levels (Fig. 4b). Notably, neuroligin protein levels were also reduced to wild-type levels after 4EGI-1 treatment (Fig. 4b).

4EGI-1 blocks long-term memory consolidation but not reconsolidation in mice[36]. In *Eif4ebp2* knockout mice a single tetanic train elicits late-phase long-term potentiation (L-LTP)[6], indicating a lower threshold for L-LTP induction compared to wild-type mice. Bath application of 4EGI-1 in acute hippocampal slices inhibited the facilitation of L-LTP induction in *Eif4ebp2* knockout mice (Supplementary Fig. 10c). Moreover, application of 4EGI-1 (Fig. 4a) to hippocampal slice cultures from *Eif4ebp2* knockout mice reversed the increase in mEPSC amplitude and frequency (Fig. 4c), reversed the increase in mIPSC amplitude (Fig. 4d), decreased the total mEPSC and mIPSC charge transfer (Fig. 4e), and restored the E/I balance.
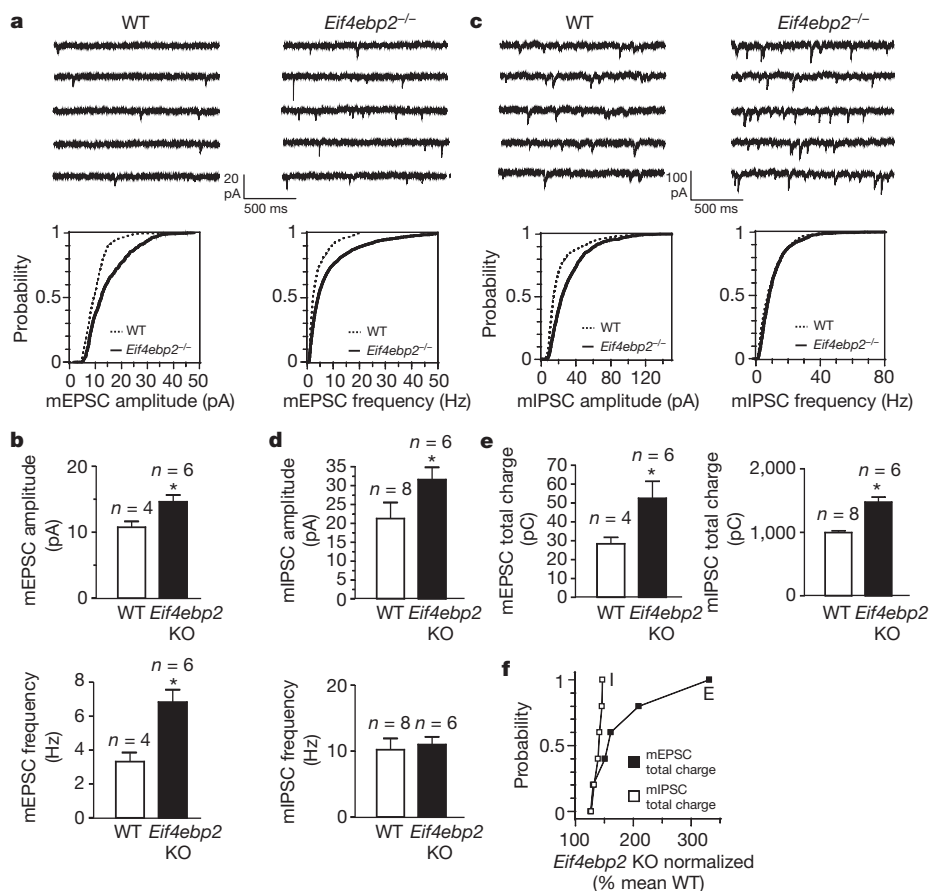
Figure 3 | Excitation is increased to a larger extent than inhibition in *Eif4ebp2* knockout mice. a, c, Representative traces (top) and cumulative distribution plots (bottom) of mEPSCs (**a**) and mIPSCs (**c**) from CA1 pyramidal neurons in acute slices from wild-type and *Eif4ebp2*$^{-/-}$ mice, illustrating changes in amplitude and frequency. **b, d**, Bar graphs showing mEPSC (**b**) and mIPSC (**d**) amplitude and frequency (*n* = number of cells).

e, Bar graphs showing the synaptic total charge transfer in *Eif4ebp2* knockout relative to wild-type slices. For **a–e**, *$P < 0.05$; Student's *t*-test. **f**, Relative changes in mEPSC and mIPSC total charge transfer, normalized to the mean wild-type value, for each neuron from *Eif4ebp2* knockout slices; Kolmogorov–Smirnov test, $P = 0.005$, KS statistic = 0.916; E, excitatory; I, inhibitory. All error bars show mean ± s.e.m.
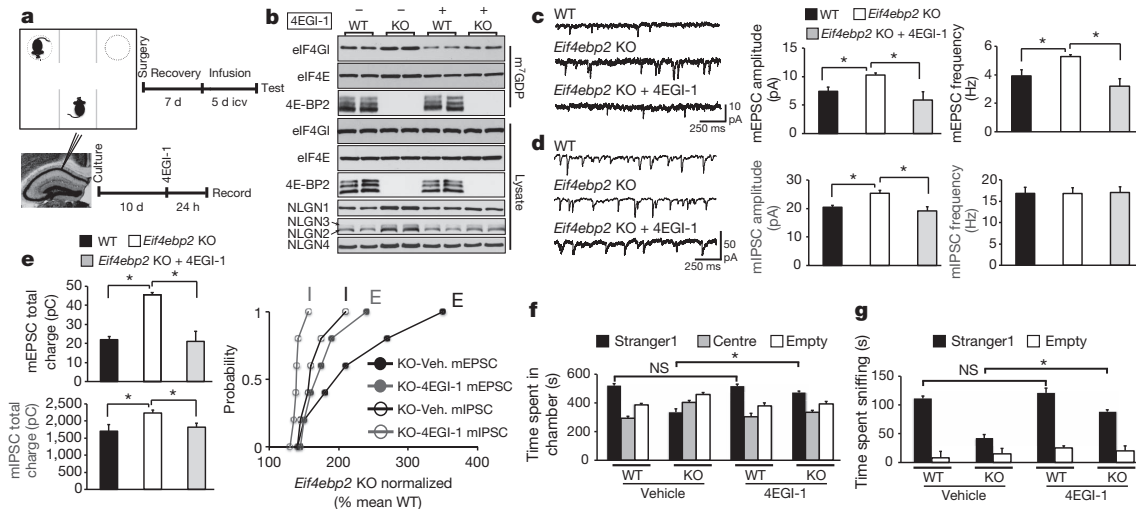
Figure 4 | Rescue of excitatory/inhibitory synaptic activity imbalance and social deficits in *Eif4ebp2* knockout mice by inhibiting the eIF4E–eIF4G interaction. **a**, Schematic of 4EGI-1 treatment in the three-chamber test (top) and slice-culture recordings (bottom). icv, intracerebroventricular infusion. **b**, Immunoblot analysis of cap column pull down in *Eif4ebp2* knockout and wild-type mice treated with 4EGI-1 or vehicle. Total and m^{7}GDP bound extracts were probed for eIF4GI, eIF4E and 4E-BP2; $n = 4$ (quantification in Supplementary Fig. 16a). **c**, **d**, Effects of 4EGI-1 on mEPSC (**c**) or mIPSC (**d**) amplitude and frequency in *Eif4ebp2* knockout slices. **e**, Effects of 4EGI-1 on total charge transfer of mEPSCs (top) and mIPSCs (bottom) in *Eif4ebp2* knockout slices. Right: relative changes in mEPSC and mIPSC total charge transfer, normalized to the mean wild-type value, for *Eif4ebp2* knockout neurons. Kolmogorov–Smirnov test. E, excitatory; I, inhibitory. For **c–e**, $n = 6$ per group; *$P < 0.05$; one-way ANOVA with Bonferroni's post-hoc test. **f**, **g**, Rescue effects of 4EGI-1 infusion in the three-chamber test showing time spent in chambers (**f**) and time spent sniffing the wire cage (**g**) in *Eif4ebp2* knockout and wild-type mice. For **f**, **g**, $n = 12$ per group; *$P < 0.03$; two-way ANOVA, with Bonferroni's post-hoc test. All error bars show mean ± s.e.m.

Next, we examined whether 4EGI-1 could rescue the social behaviour deficits in *Eif4ebp2* knockout mice. We used intracerebroventricular cannulation to infuse 4EGI-1 daily for 5 days (Fig. 4a). Mice were then subjected to the three-chamber social interaction test. Whereas a sub-threshold concentration of 4EGI-1 had no effect on wild-type mice (Fig. 4f and Supplementary Fig. 10b), it reversed social interaction deficits of *Eif4ebp2* knockout mice, inasmuch as 4EGI-1-treated *Eif4ebp2* knockout mice spent significantly more time in the stranger1 chamber (Fig. 4f), whereas the number of entries for the *Eif4ebp2* knockout mice infused with 4EGI-1 was restored to wild-type levels (Supplementary Fig. 12a). Also, the interaction time with stranger1 was increased compared to vehicle-treated knockout mice (Fig. 4g). These results establish causality between inactivation of the eIF4F complex and rescue of the social interaction impairments in the *Eif4ebp2* knockout mice.

Nlgn1 knockdown rescues ASD-like *Eif4ebp2* phenotypes

NLGN1, which is present exclusively at excitatory synapses, promotes excitatory synaptic transmission[37,38]. In *Eif4ebp2* knockout mice *Nlgn1* mRNA translation is enhanced concomitant with increased E/I ratio (Figs 2 and 3). We therefore predicted that reduction of NLGN1 levels in *Eif4ebp2* knockout mice would restore the E/I balance and reverse the autism-related deficits. To investigate this, we used short interfering RNAs (siRNAs) against *Nlgn1* in hippocampal pyramidal cells in slice cultures (Fig. 5a). Knockdown of *Nlgn1* rescued the amplitude and frequency of mEPSCs in *Eif4ebp2* knockout pyramidal cells (Fig. 5c), without affecting mIPSCs (Fig. 5d). As expected, mEPSC total charge transfer in pyramidal cells of *Eif4ebp2* knockout mice transfected with *Nlgn1* siRNA was restored to wild-type levels (Fig. 5e), whereas mIPSC total charge transfer was unaffected (Fig. 5e). Next, we knocked down *Nlgn2*, a regulator of inhibitory synapses. In contrast to NLGN1, knockdown of *Nlgn2* rescued the increase in mIPSC amplitude in *Eif4ebp2* knockout pyramidal cells (Fig. 5d), without affecting the amplitude or frequency of mEPSCs (Fig. 5c). The increase in total charge transfer of mIPSCs was also attenuated in *Nlgn2*-siRNA-transfected *Eif4ebp2* knockout pyramidal cells, whereas mEPSC total charge transfer remained unaltered (Fig. 5e).

We next subjected wild-type or *Eif4ebp2* knockout mice to the three-chamber social interaction test after intracerebroventricular injection of shRNAs (Fig. 5a and Supplementary Fig. 11). Notably, NLGN1 or NLGN2 protein levels decreased to wild-type levels in *Eif4ebp2* knockout mice injected with *Nlgn1* or *Nlgn2* shRNA, respectively, whereas protein levels of the other neuroligins were not changed (Fig. 5b). We selected a concentration of lentiviral shRNA (Supplementary Fig. 11) that had no effect in wild-type mice (Fig. 5f, g). Importantly, this sub-threshold concentration of *Nlgn1* shRNA partially rescued the social interaction impairment in *Eif4ebp2* knockout mice as follows: *Eif4ebp2* knockout mice injected with *Nlgn1* shRNA spent significantly more time in the stranger1 chamber (Fig. 5f), entered the chamber more times (Supplementary Fig. 12b), and interacted more with stranger1 (Fig. 5g) relative to control knockout mice. In contrast, knockdown of *Nlgn2* exacerbated the social deficits in *Eif4ebp2* knockout mice, as shown by the reduced time spent in the stranger1 chamber (Fig. 5f), and the decreased amount of time spent interacting with stranger1 (Fig. 5g) as compared to control knockout mice. We recapitulated these findings with additional shRNAs against *Nlgn1* or *Nlgn2* (Supplementary Fig. 13). In summary, NLGN1 knockdown reverses changes at excitatory synapses and partially rescues the social interaction deficits in *Eif4ebp2* knockout mice, whereas NLGN2 knockdown weakens inhibitory synapses and exacerbates the phenotype, thus establishing a strong link between eIF4E-dependent translational control of neuroligins, E/I balance and the development of ASD-like phenotypes.

Discussion

We show that 4E-BP2 deficiency in mice causes autistic-like phenotypes that markedly resemble behavioural hallmarks of ASD. *Eif4ebp2* knockout mice display a previously undocumented deficit in social interactions and exhibit repetitive/stereotyped behaviours. Communication between *Eif4ebp2* knockout pups and their dams is altered, as demonstrated by enhanced ultrasonic vocalizations. This phenotype is similar to that reported for *Tsc2*[+/−] (ref. 13) or conditional *Tsc2* knockout[14] mice, which could conceivably be explained by increased eIF4E activity. These newly discovered autistic-like phenotypes of *Eif4ebp2* knockout mice are neither confounded by anxiety
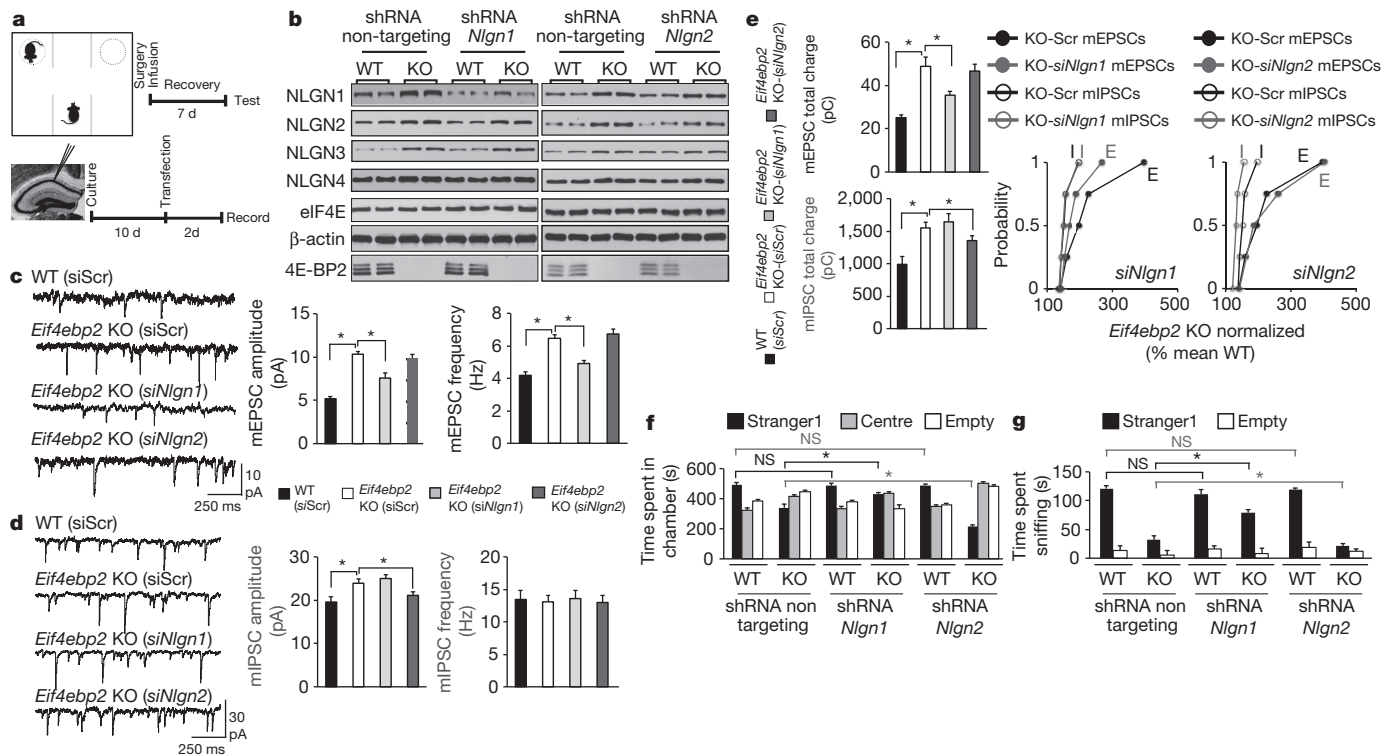
Figure 5 | Knockdown of neuroligin 1 rescues the excitatory/inhibitory synaptic activity imbalance and social deficits in *Eif4ebp2* knockout mice. a, Schematic of the three-chamber test in mice intracerebroventricularly injected with lentiviruses expressing shRNAs (top) and recordings from siRNA-transfected slice cultures (bottom). **b**, Western blot analysis of hippocampal lysates from wild-type or *Eif4ebp2* knockout mice injected with lentiviruses expressing shRNAs against a non-targeting sequence or *Nlgn1* or *Nlgn2*. Extracts were probed for NLGN1– NLGN4, β-actin, eIF4E and 4E-BP2; $n = 4$ (quantification in Supplementary Fig. 16b, c). **c, d**, Effects of *Nlgn1* or *Nlgn2* knockdown on mEPSC (**c**) or mIPSC (**d**) amplitude and frequency in

transfected pyramidal cells. *siNlgn1* or *siNlgn2*, siRNA against *Nlgn1* or *Nlgn2*; siScr, scrambled siRNA. **e**, mEPSC (top) or mIPSC (bottom) total charge transfer in *Nlgn1* or *Nlgn2* knockdown cells. Right: relative changes in mEPSC and mIPSC total charge transfer, normalized to the mean wild-type value, for each neuron from *Eif4ebp2* knockout slices; Kolmogorov–Smirnov test. For **c–e**, $n = 5$ per group; $*P < 0.05$; one-way ANOVA with Bonferroni's post-hoc test. **f, g**, Knockdown of *Nlgn1* or *Nlgn2* in the three-chamber test showing time spent in chambers (**f**) and time spent sniffing the wire cage (**g**). For **f, g**, $n = 12$ per group; $*P < 0.02$; two-way ANOVA, with Bonferroni's post-hoc test. All error bars show mean ± s.e.m.

nor by the previously reported long-term memory deficits[6,39], as short-term memory circuits are intact. Thus, *Eif4ebp2* knockout mice constitute a novel model for mTORC1/eIF4E-dependent autism-like phenotypes due to dysregulated translational control. In accordance with this model, a point mutation in the promoter region of the *EIF4E* gene is associated with increased activity of eIF4E in autistic siblings from two unrelated families[40].

In two models of enhanced eIF4E activity (*Eif4ebp2* knockout and eIF4E-overexpression mice), we show that neuroligin mRNAs are translated more efficiently. These results highlight the important role of 4E-BP2 and eIF4E in postsynaptic translational control of neuronal mRNAs. Increased mTOR signalling to its downstream effectors promotes the preferential translation of a subset of mRNAs referred to as 'eIF4E sensitive', harbouring extensive secondary structure elements in their 5′ UTRs[5,41]. Reporter mRNAs containing the full-length 5′ UTRs of *Nlgn1* and *Nlgn2*, but not *Dlg4* or *Actb*, fused to luciferase, are better translated in cells with elevated mTORC1 signalling (*Pten*[+/−], *Tsc2* knockout) and/or enhanced cap-dependent translation (*Eif4ebp2* knockout, eIF4E overexpression) as compared to parental cells (Supplementary Fig. 14). Interestingly, the 5′ UTRs of neuroligin mRNAs possess a repeated structural element that is absent from the other mRNA 5′ UTRs (Supplementary Fig. 15). These secondary structure repeats might account for the preferential increase in translation of neuroligins in response to increased eIF4F activity.

Mutations and copy number variation in neuroligins have been associated with ASD[42]. In *Eif4ebp2* knockout mice, increased amounts of NLGN1– NLGN4 lead to a disproportionate ratio of neuroligins to

PSD95 (in excitatory synapses) or to gephyrin (in inhibitory synapses) because the general composition of the postsynaptic scaffold (other than neuroligins) is not altered by *Eif4ebp2* deletion or similarly over-expression of eIF4E. An altered E/I synaptic ratio is a prominent feature in ASD and in mouse models of autism[19,21]. The relative impact of *Eif4ebp2* on excitation is greater than on inhibition, thus engendering a shift in the E/I balance. In *Eif4ebp2* knockout mice, pharmacological inhibition of the eIF4E–eIF4G interaction using the selective inhibitor 4EGI-1 reduces eIF4F activity to wild-type levels, restores the L-LTP threshold and rescues the E/I imbalance phenotype. In *Eif4ebp2* knockout mice, knockdown of *Nlgn1* restores the E/I balance and improves social approach behaviour to wild-type levels. In contrast, *Nlgn2* knockdown increases the E/I imbalance and exacerbates the social behaviour deficits. These results establish a link between eIF4E translational control of neuroligins, E/I balance and ASD-like phenotypes (Supplementary Fig. 1).

This novel regulatory mechanism should provide an impetus to investigate the role of eIF4E-mediated translation in ASD further and to explore therapeutic avenues for the pharmacological modulation of eIF4F activity (for example, with 4EGI-1 and its derivatives) or the function of neuroligins.

METHODS SUMMARY

All procedures are in accordance with the collaborating institutions' Animal Care Committee guidelines. Experimenters were blinded to the genotype during testing and scoring. All data are presented as mean ± s.e.m. (error bars). Statistical results, along with tests used, are summarized in Supplementary Table 4. Statistical significance was set a priori at 0.05.

1. Fombonne, E. Epidemiology of pervasive developmental disorders. *Pediatr. Res.* **65,** 591–598 (2009).
2. Kelleher, R. J. & Bear, M. F. The autistic neuron: troubled translation? *Cell* **135,** 401–406 (2008).
3. Hay, N. & Sonenberg, N. Upstream and downstream of mTOR. *Genes Dev.* **18,** 1926–1945 (2004).
4. Pause, A. *et al.* Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5′-cap function. *Nature* **371,** 762–767 (1994).
5. Koromilas, A. E., Lazaris-Karatzas, A. & Sonenberg, N. mRNAs containing extensive secondary structure in their 5′ non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E. *EMBO J.* **11,** 4153–4158 (1992).
6. Banko, J. L. *et al.* The translation repressor 4E–BP2 is critical for eIF4F complex formation, synaptic plasticity, and memory in the hippocampus. *J. Neurosci.* **25,** 9581–9590 (2005).
7. Zhou, J. & Parada, L. F. PTEN signaling in autism spectrum disorders. *Curr. Opin. Neurobiol.* http://dx.doi.org/10.1016/j.conb.2012.05.004 (2012).
8. Kwon, C. H. *et al.* Pten regulates neuronal arborization and social interaction in mice. *Neuron* **50,** 377–388 (2006).
9. Zhou, J. *et al.* Pharmacological inhibition of mTORC1 suppresses anatomical, cellular, and behavioral abnormalities in neural-specific Pten knock-out mice. *J. Neurosci.* **29,** 1773–1783 (2009).
10. Jeste, S. S., Sahin, M., Bolton, P., Ploubidis, G. B. & Humphrey, A. Characterization of autism in young children with tuberous sclerosis complex. *J. Child Neurol.* **23,** 520–525 (2008).
11. Auerbach, B. D., Osterweil, E. K. & Bear, M. F. Mutations causing syndromic autism define an axis of synaptic pathophysiology. *Nature* **480,** 63–68 (2011).
12. Ehninger, D. *et al.* Reversal of learning deficits in a *Tsc2*^+/− mouse model of tuberous sclerosis. *Nature Med.* **14,** 843–848 (2008).
13. Young, D. M., Schenk, A. K., Yang, S. B., Jan, Y. N. & Jan, L. Y. Altered ultrasonic vocalizations in a tuberous sclerosis mouse model of autism. *Proc. Natl Acad. Sci. USA* **107,** 11074–11079 (2010).
14. Tsai, P. T. *et al.* Autistic-like behaviour and cerebellar dysfunction in Purkinje cell *Tsc1* mutant mice. *Nature* **488,** 647–651 (2012).
15. O'Roak, B. J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of *de novo* mutations. *Nature* **485,** 246–250 (2012).
16. Hoeffer, C. A. *et al.* Altered mTOR signaling and enhanced CYFIP2 expression levels in subjects with fragile X syndrome. *Genes Brain Behav.* **11,** 332–341 (2012).
17. Napoli, I. *et al.* The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP. *Cell* **134,** 1042–1054 (2008).
18. Nowicki, S. T. *et al.* The Prader-Willi phenotype of fragile X syndrome. *J. Dev. Behav. Pediatr.* **28,** 133–138 (2007).
19. Rubenstein, J. L. & Merzenich, M. M. Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes Brain Behav.* **2,** 255–267 (2003).
20. Uhlhaas, P. J. & Singer, W. Neuronal dynamics and neuropsychiatric disorders: toward a translational paradigm for dysfunctional large-scale networks. *Neuron* **75,** 963–980 (2012).
21. Cornew, L., Roberts, T. P., Blaskey, L. & Edgar, J. C. Resting-state oscillatory activity in autism spectrum disorders. *J. Autism Dev. Disord.* **42,** 1884–1894 (2012).
22. Yizhar, O. *et al.* Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* **477,** 171–178 (2011).
23. Luikart, B. W. *et al.* Pten knockdown *in vivo* increases excitatory drive onto dentate granule cells. *J. Neurosci.* **31,** 4345–4354 (2011).
24. Bateup, H. S., Takasaki, K. T., Saulnier, J. L., Denefrio, C. L. & Sabatini, B. L. Loss of Tsc1 *in vivo* impairs hippocampal mGluR-LTD and increases excitatory synaptic function. *J. Neurosci.* **31,** 8862–8869 (2011).
25. Peça, J. *et al.* Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature* **472,** 437–442 (2011).
26. Schmeisser, M. J. *et al.* Autistic-like behaviours and hyperactivity in mice lacking ProSAP1/Shank2. *Nature* **486,** 256–260 (2012).
27. Won, H. *et al.* Autistic-like social behaviour in Shank2-mutant mice improved by restoring NMDA receptor function. *Nature* **486,** 261–265 (2012).
28. Chao, H. T. *et al.* Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. *Nature* **468,** 263–269 (2010).
29. Levinson, J. N. & El-Husseini, A. Building excitatory and inhibitory synapses: balancing neuroligin partnerships. *Neuron* **48,** 171–174 (2005).
30. Scattoni, M. L., Crawley, J. & Ricceri, L. Ultrasonic vocalizations: a tool for behavioural phenotyping of mouse models of neurodevelopmental disorders. *Neurosci. Biobehav. Rev.* **33,** 508–515 (2009).
31. Ruggero, D. *et al.* The translation factor eIF-4E promotes tumor formation and cooperates with c-Myc in lymphomagenesis. *Nature Med.* **10,** 484–486 (2004).
32. Graff, J. R. *et al.* Therapeutic suppression of translation initiation factor eIF4E expression reduces tumor growth without toxicity. *J. Clin. Invest.* **117,** 2638–2648 (2007).
33. Glessner, J. T. *et al.* Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* **459,** 569–573 (2009).
34. Wang, K. *et al.* Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* **459,** 528–533 (2009).
35. Moerke, N. J. *et al.* Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. *Cell* **128,** 257–267 (2007).
36. Hoeffer, C. A. *et al.* Inhibition of the interactions between eukaryotic initiation factors 4E and 4G impairs long-term associative memory consolidation but not reconsolidation. *Proc. Natl Acad. Sci. USA* **108,** 3383–3388 (2011).
37. Chubykin, A. A. *et al.* Activity-dependent validation of excitatory versus inhibitory synapses by neuroligin-1 versus neuroligin-2. *Neuron* **54,** 919–931 (2007).
38. Dahlhaus, R. *et al.* Overexpression of the cell adhesion protein neuroligin-1 induces learning deficits and impairs synaptic plasticity by altering the ratio of excitation to inhibition in the hippocampus. *Hippocampus* **20,** 305–322 (2010).
39. Banko, J. L. *et al.* Behavioral alterations in mice lacking the translation repressor 4E–BP2. *Neurobiol. Learn. Mem.* **87,** 248–256 (2007).
40. Neves-Pereira, M. *et al.* Deregulation of EIF4E: a novel mechanism for autism. *J. Med. Genet.* **46,** 759–765 (2009).
41. Larsson, O. *et al.* Distinct perturbation of the translatome by the antidiabetic drug metformin. *Proc. Natl Acad. Sci. USA* **109,** 8977–8982 (2012).
42. Levy, D. *et al.* Rare *de novo* and transmitted copy-number variation in autistic spectrum disorders. *Neuron* **70,** 886–897 (2011).

METHODS

Knockout and knock-in mice. *Eif4ebp2* knockout and βT-*Eif4e* mice were previously described[10,34]. Where mentioned, wild-type mice are littermates of the corresponding genotype knockout or knock-in mice. For all experiments, littermates from heterozygote crossings were used. Mice were backcrossed for more than ten generations to C57BL/6J mice. Food and water were provided ad libitum, and mice were kept on a 12-h light/dark cycle. Pups were kept with their dams until weaning at postnatal day 21. After weaning, mice were group housed (maximum of four per cage) by sex, or housed in pairs (for home cage social interaction). Cages were maintained in ventilated racks in temperature (20–21 °C) and humidity (~55%) controlled rooms, on a 12-h circadian cycle (7:00–19:00 light period). Standard bedding was used for housing with the addition of Enviro-Dri (Shepherd Speciality Paper) and a small cardboard house. All procedures were in compliance with the Canadian Council on Animal Care guidelines and were approved by McGill University, UCSF and Université de Montréal Animal Care Committees. For all behavioural assays, mice were handled for three consecutive days before experimental testing.

Three-chamber social approach and preference for social novelty tests. A three-chamber arena was used to assess social approach and preference for social novelty[43]. Test mice (2-month-old males) were placed in the middle chamber and allowed to explore all the empty chambers of the apparatus freely for 10 min. Immediately after habituation, an unfamiliar mouse (stranger1, male C57BL/6 age-matched) was introduced into one of the two side chambers, enclosed in a wire cage allowing only the test mouse to initiate any social interaction. An identical empty wire cage was placed in the other side chamber. After placement, the side doors were opened simultaneously and the test mouse was allowed to explore the whole three-chamber arena for 20 min. At the end of the 20-min sociability test, a new unfamiliar mouse (stranger2, male C57BL/6 age-matched) was placed in the previously unoccupied wire cage, and test mice were examined for an additional 20 min to assess preference for social novelty. The location of the empty wire cage was alternated between side chambers for different subjects. The time spent in each compartment, the number of transitions between chambers and the time spent sniffing the stranger1, stranger2 or empty wire cages, were manually scored.

Home cage social interaction test. Wild-type or knockout *Eif4ebp2* male littermate mice were weaned and housed in pairs of wild-type–wild-type, knockout–knockout or wild-type–knockout. At the age of 2 months, mice were recorded for 20 min to assess home cage social interactions. The cage lid, grid and cardboard nest were removed for the recording session, and time in close contact (huddling, sniffing (nose-to-nose, anogenital sniffing) and allogrooming) was scored for each mouse. A small mark with an odourless, permanent marker and an ear punch (both marked the day before the test), were used to identify each genotype. Tagging was alternated randomly between wild-type and knockout mice in the various sessions.

Reciprocal social interaction test. Subjects were unfamiliar *Eif4ebp2* knockout or wild-type mice originating from separate litters. The social interaction in pairs (wild-type–wild-type, wild-type–knockout or knockout–knockout) test was performed in a neutral cage (grey Plexiglas box, 30 × 30 × 30 cm). Mice were first habituated for 2 days in 10-min sessions and on day 3 pairs of unfamiliar mice with the same genotype were placed into the neutral cage for 10 min. The time spent in close contact (indicative of social interaction) was registered. Close contact is defined as close huddling, sniffing (nose-to-nose, anogenital sniffing) or allogrooming by the target and stimulus mouse. A small mark with an odourless, permanent marker and an ear punch (both marked the day before the test) were used to identify each genotype. Tagging was alternated randomly between wild-type and knockout mice in the various sessions.

Self-grooming test. Wild-type or *Eif4ebp2* knockout mice were placed in a new Plexiglas cage with fresh bedding and no nest or cardboard material. Self-grooming behaviour was recorded for 10 min after an initial 10-min habituation phase.

Marble-burying task. Mice were individually placed in Plexiglas cages containing 5-cm-deep fresh bedding with 20 black marbles pre-arranged in 5 × 4 evenly spaced rows[44]. Testing was conducted for 20 min. After the test period, unburied marbles were counted. Marbles were considered buried if they were at least one-half covered with bedding.

Isolation-induced USVs in mouse pups. A microphone (Knowles Acoustics FG-3329C) was fixed inside the top of an anechoic glass chamber wrapped with a heated mat maintaining a constant temperature of 21 °C. The microphone was connected to a bat detector (Ultrasonic Detector D 230; Pettersson Elktronic AB), which was set for frequency division mode so that the entire frequency range (10–120 kHz) was divided by 10 and thus made audible to the human ear. The frequency division system of the D 230 retains the original input amplitude. The signal from the bat detector was amplified through a MIDI interface (US-122; Tascam TEAC Professional Division) and fed to a standard PC for recording

using a digital recording system (MITSYN, WLH, Version 2005.7.0.0). Newborn pups (postnatal days 2, 4, 6, 8 and 12) were carefully separated from their dams and placed individually in the recording chamber. USVs were recorded for 5 min. Recordings were analysed blind to genotype with MITSYN (audio-editing) software.

Elevated plus maze test. For the elevated plus maze apparatus we constructed a cruciform platform consisting of: two open arms (30 cm × 5 cm) assembled opposite each other and two enclosed arms (30 cm × 5 cm × 15 cm) also facing each other. A centre platform (5 cm × 5 cm) was used as an anchor point for all arms and the entire apparatus was elevated 40 cm off the floor. Low light and white noise conditions were used in the test room. Each mouse was placed in the centre of the maze facing an open arm. During the testing period of 5 min, the total time spent in each arm was recorded.

Contextual fear conditioning test. Wild-type or *Eif4ebp2* knockout mice were cannulated, as described in the stereotaxic surgery section. During acquisition, two foot shocks of 0.5 mA for 2 s separated by 60 s were administered after the 2-min period of context exploration. Immediately after acquisition, mice were infused with different concentrations of 4EGI-1. Twenty-four hours after training, mice were tested for contextual fear memory, as assessed by percentage freezing in the conditioning context for a 5-min period, in 5-s intervals, either 'freezing' or 'not freezing'. Freezing (%) indicates the number of intervals where freezing was observed divided by total number of 5-s intervals.

Polysome profile analysis. Polysome profile analysis was carried out as previously described[41]. Intact hippocampi were washed with ice-cold PBS containing 100 µg ml^{-1} cycloheximide and subsequently lysed in a hypotonic lysis buffer (5 mM Tris-HCl (pH 7.5), 2.5 mM MgCl$_2$, 1.5 mM KCl, 100 µg ml^{-1} cycloheximide, 2 mM DTT, 0.5% Triton X-100, and 0.5% sodium deoxycholate). Lysate concentration was balanced using a Bradford-assay (BIORAD) and by measuring total RNA concentration using a NANODROP2000 spectrophotometer (Thermo Scientific). Lysates were loaded onto 10–50% sucrose density gradients (20 mM HEPES-KOH (pH 7.6), 100 mM KCl, 5 mM MgCl$_2$) and centrifuged at 35,000 r.p.m. for 2 h at 4 °C. Gradients were fractionated and the absorbance at 254 nm was continuously recorded using an ISCO fractionator (Teledyne ISCO). Total RNA from each fraction was isolated using Trizol (Invitrogen) and reverse transcribed using the Superscript III kit (Invitrogen). Polysome to monosome ratio was calculated as the area under the A_{254} absorbance curve, using the function describing the absorbance values, processed with the definite integral command in MATLAB.

qRT–PCR and RT–PCR. Total or polysome reverse-transcribed RNA was analysed using a Biorad iQ SYBR Green Supermix kit as previously described[41]. For all experiments, $n = 4$ and the primers used are summarized in Supplementary Table 3. For RT–PCR results depicted on agarose gels, products were amplified for 20–25 cycles, to remain within the linear range of the SYBR green reaction. Results are presented in arbitrary units as relative amounts using serial dilutions of cortical or hippocampal RNA as qRT–PCR concentration standards.

Western blotting and antibodies. All tissues and cells were dissociated in RIPA buffer (unless otherwise specified). Western blotting was previously described[45]. Antibodies against indicated proteins were: neuroligin 1 and neuroligin 2/3, neuroligin 2, neuroligin 3, VGAT, VGLUT (Synaptic Systems); neuroligin 4 (Abcam); PSD95, GFAP, 4E-BP2, β-actin, eIF4E and eIF4GI (Cell Signaling); neurexin1 (α, β), neurexin2 (α, β), neurexin3 (α, β), SAPAP3, SHANK2, SHANK3, gephyrin, Arc/Arg3.1 and POLR2A (Santa-Cruz); β-tubulin (Sigma); secondary anti-mouse and anti-rabbit (GE Healthcare); and anti-goat (Santa-Cruz). Quantification of immunoblots was performed using ImageJ (NIH). Values were normalized to β-actin or another control where specified.

Synaptosome preparation. Intact hippocampi from wild-type, *Eif4ebp2* or βT-*Eif4e* mice were dissected in ice-cold PBS and homogenized in 320 mM sucrose, 1 mM EDTA, 5 mM Tris-HCl (pH 7.4) and 25 µM DTT. Synaptosomes were isolated on a discontinuous Percoll (GE Healthcare) gradient. The fraction between 5% and 23% Percoll was isolated and re-suspended in 2× SDS–PAGE sample buffer. POLR2A (RNA polymerase II large subunit) is used as a nuclear marker, GFAP (glial fibrillary acidic protein) as a glial marker, and Arc/Arg3.1 (activity-regulated cytoskeleton-associated protein) as a marker for synaptosomes.

m^{7}GDP pull-down. Dorsal hippocampi were dissected in cold PBS, and homogenized in lysis buffer containing 40 mM HEPES-KOH (pH 7.5), 120 mM NaCl, 1 mM EDTA, 0.1 mM GDP, 10 mM pyrophosphate, 10 mM β-glycerophosphate, 50 mM NaF and 0.3% CHAPS. Extracts were mixed with 30 µl of m^{7}GDP-agarose (GE Healthcare) for 1.5 h at 4 °C. The resins were washed four times and proteins were eluted with 2× SDS–PAGE sample buffer.

Golgi staining and immunostaining. Hippocampal slices (50 µm) were prepared from whole brains of 2-month-old mice after fixation (4% PFA). Golgi staining of neuronal somata and processes was carried out using the FD Rapid Golgi Stain Kit (FD NeuroTechnologies) following the manufacturer's protocol. Images of dendritic spines from hippocampal CA1 neurons were acquired using a

×63 oil lens on a Zeiss Axioscope. For immunostaining, 2-month-old mice were perfused with 4% PFA, intact brains were removed and hippocampal slices were prepared. Slices were briefly permeabilized with 50 mM NH_4Cl and 0.2% (w/v) saponin (SIGMA) in phosphate buffered saline (PBS), blocked in 5% goat serum, incubated with primary antibodies (1:200) for 3 h, washed with PBS, incubated with Alexa Fluor 488 or 594 (Molecular Probes), washed with PBS and mounted on microscope coverslips with Vectashield mounting medium (Vector Labs). Images were acquired with a Zeiss confocal microscope using a ×63 oil-immersion lens.

Field EPSP recordings. Transverse hippocampal slices (400 μm) were prepared from 6-week-old wild-type or *Eif4ebp2* knockout littermates as described earlier[45]. Briefly, slices were submerged and superfused with an oxygenated artificial cerebrospinal fluid (ACSF) containing: 124 mM NaCl, 2.5 mM KCl, 1.25 mM NaH_2PO_4, 1.3 mM $MgSO_4$, 2.5 mM $CaCl_2$, 26 mM $NaHCO_3$ and 10 mM glucose at 28 °C. Slices were allowed to recover at least 2 h before recording was initiated. A concentric bipolar tungsten stimulating electrode placed in the middle of CA1 stratum radiatum was used to stimulate Schaffer collateral afferents and elicit field extracellular field potentials (fEPSPs). fEPSPs were recorded with ASCF-filled glass electrode (2–3 MΩ) placed in CA1 stratum radiatum ~200 mm from the stimulating electrode. Baseline stimulation was applied at 0.033 Hz by delivering 0.1-ms pulses that evoked 35% of maximal fEPSPs. LTP was induced with a single train at 100 Hz for 1 s. 4EGI-1 (50 μM, Calbiochem) was applied 30 min before the onset of tetanic stimulation and for 1 h.

Miniature EPSC and IPSC recordings. For whole-cell recordings, acute slices were prepared as described above but from younger mice (19–22 days old). Organotypic hippocampal slice cultures were prepared and biolistically transfected as previously described[45]. siRNAs against *Nlgn1*, *Nlgn2* or scrambled (*Nlgn1*: 5′-UCUGGGAAUAACUGGUCCAUGCUACUU-3′, *Nlgn2*: 5′-GUC CUCACUCUCUCCAGAGUCCUCCAC-3′ or scrambled: NC1 negative control) were designed using siRNA-design Software (IDT). An enhanced yellow fluorescent protein (EYFP)-expressing plasmid was co-transfected with all siRNAs (pcDNA3.1-EYFP). Whole-cell recordings were obtained from CA1 pyramidal neurons using borosilicate pipettes (3–6 MΩ). The intracellular solution for recording EPSCs contained: 130 mM CsMeSO3, 5 mM CsCl, 2 mM $MgCl_2$, 5 mM diNa-phosphocreatine, 10 mM HEPES, 2 mM ATP-Tris, 0.4 mM GTP-Tris, and 0.1% biocytin (pH 7.2–7.3; 275–285 mOsmol). Recordings were made in the voltage-clamp mode using a Multiclamp 700A amplifier (Molecular Devices). Holding potential was maintained at −60 mV and series resistance was routinely monitored. Recorded signals were low-pass-filtered at 2 kHz, digitized at 20 kHz and stored on a PC. Data acquisition and off-line analyses were performed using 1322A Digidata acquisition board, and pClamp 10 software (Molecular Devices). Data were only included if the holding current was stable or if series resistance varied <25% of initial value. Miniature EPSCs (mEPSCs) or IPSCs (mIPSCs) were recorded in the presence of TTX (0.5 μM; Alomone Labs). AMPA receptor-mediated mEPSCs were isolated by co-application of DL-2-amino-5-phosphonovaleric acid (AP5; 50 μM; Tocris Biosciences) and gabazine (5 μM; Tocris Biosciences), to block NMDA and $GABA_A$ receptors, respectively. $GABA_A$ receptor-mediated mIPSCs were recorded using a CsCl-based intracellular solution which contained: 135 mM CsCl, 10 mM HEPES, 2 mM $MgCl_2$ and 4 mM MgATP (pH 7.25). In these experiments, 6,7-dinitroquinoxaline-2,3-dione (DNQX; 5 mM) and AP5 (50 mM) were present to block excitatory transmission via AMPA and NMDA receptors. For analysis, mEPSCs or mIPSCs were detected on a running template (mean of ~ 20 events) with a well-defined baseline, using pClamp10 software. Detection threshold was set at 3 pA and 150–200 events were sampled per neuron. Total charge was calculated by summing the charge transfer of all individual events (mEPSCs or mIPSCs) detected over a 10-min acquisition period for each neuron.

Evoked EPSCs and IPSCs. Putative single-fibre EPSCs or IPSCs were evoked using minimal stimulation. Stimulus strength was adjusted to a value that generated ~ 50% successes (range 40–60%). Amplitudes of mean EPSC or IPSC were extracted from 5-min bins over a 10–20-min period.

Metabolic labelling. Acute hippocampal slices were prepared from wild-type or *Eif4ebp2* knockout or βT-*Eif4e* mice as described above and 30 mCi ml⁻¹ [35S]Met/Cys was added for 3 h. Slices were homogenized in 25 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1 mM EDTA, 1% NP-40 and 5% glycerol, and lysates were centrifuged at 16,000g. Total extracts were resolved with SDS–PAGE electrophoresis and transferred onto nitrocellulose membranes, which were then exposed onto autoradiography film (KODAK) (1 week) and developed. For total protein calculations, lysates were analysed with SDS–PAGE electrophoresis and gels stained with GelCode blue Stain Reagent (Pierce) according to the manufacturer's protocol. Signals were quantified using ImageJ.

Cloning of 5′ UTR luciferase constructs. Full-length 5′ UTRs of mouse *Nlgn1* (NCBI reference NM_001163387.1), *Nlgn2* (NM_198862.2), *Actb* (β-actin; NM_007393.3) and *Dlg4* (PSD95) (NM_007864.3) were amplified from mouse genomic DNA, using specific primers and cloned into pGL4.13 firefly luciferase vector (Promega) upstream of the FF luciferase gene.

MEF and transient transfection. *Eif4ebp2* knockout, *Eif4e* overexpressing, *Tsc2* knockout and *Pten*⁺/⁻ mouse embryonic fibroblasts (MEFs) were prepared as previously described[46]. Cells were maintained in DMEM medium (Wisent) supplemented with 10% FBS (Wisent) and 5% penicillin/streptomycin (Wisent) in a humidified, 37 °C, 5% CO_2 incubator. Transient transfections were carried out for 3 h with lipofectamine 2000 (Invitrogen) as per the manufacturer's recommendations (using a 3:1 ratio of Firefly-expressing pGL4.13 to Renilla expressing vectors). Renilla and firefly luciferase signals were detected 24 h after transfection with a STOPnGlo kit (Promega), using a Fluostar OPTIMA luminometer (BMG Labtech).

Stereotaxic surgery, cannulation and infusion. Two-month-old mice were deeply anaesthetized with 1 ml kg⁻¹ of body weight of a mixture of ketamine (55 mg per kg), xylazine (3 mg per kg) and medetomidine hydrochloride (0.3 mg per kg) injected intraperitoneally. Mice were placed in a Kopf stereotactic frame. A small hole was made 1 mm rostral of the bregma and 2 mm right of midline and stainless steel 22-gauge cannulas were implanted at these coordinates. Obturators inserted in the guides prevented blocking. Dental cement was applied to stabilize the implants. After surgery, an intramuscular injection of analgesic (carprofen, 5 mg per kg) was given. An intraperitoneal injection of Antisedan (0.3 mg per kg) suspended anaesthesia. Mice were then allowed to recover from surgery for 7 days and were handled daily during the recovery period. Infusions were carried out daily for 5 days at a depth of 2 mm. 4EGI-1 (Calbiochem) was prepared in 50% DMSO in 100 mM Tris-Cl (pH 7.2) and infused daily at a rate of 0.1 μl min⁻¹ at a final concentration of 20 μg μl⁻¹; this sub-threshold concentration had no effect in wild-type mice. At the end of the experiment cannula placement and injection site accuracy were checked by infusion of 1 μl of eosin Y (Sigma). For lentivirus injections, the same coordinates were used for a single injection without cannulas. A total of 3 μl of lentiviral particles at 1.4×10^7 TU (transducing units) ml⁻¹ were injected. Mice were returned to their home cage for 7 days (and handled during the last 3) before behavioural testing.

Lentiviral transduction and titration. Lentiviruses were produced by co-transfection of HEK293T cells with transfer vector (pLKO.1), packaging plasmid psPAX2 and envelope plasmid pMD2.G as described in http://lentiweb.com/protocols_lentivectors.php. For mouse *Nlgn1* knockdown, shRNA1 was TRCN0000032022, shRNA2 was TRCN0000032020, mouse *Nlgn2* knockdown shRNA1 was TRCN0000180497 and shRNA2 was TRCN0000184441, the non-targeting shRNA was SHC002 (all constructs from Mission Sigma). N2A cells used for validation of shRNAs and siRNAs or MEFs used for titration were maintained in DMEM (Gibco) supplemented with 10% FBS and 1% penicillin/streptomycin (Gibco) at 37 °C, 5% CO_2. Viral titre (TU ml⁻¹) was calculated using puromycin selection of MEFs infected with different viral particle dilutions and stained with 0.2% crystal violet and 20% methanol solution. Colonies were measured using the CellCount plugin of ImageJ (NIH). Viral titres were adjusted to 1.4×10^7 TU ml⁻¹ for *in vivo* injections.

Statistical analysis and behavioural scoring. Experimenters were blinded to the genotype during testing and scoring. All data are presented as mean ± s.e.m. (error bars). Statistical results, along with tests used, are summarized in Supplementary Table 4. Statistical significance was set a priori at 0.05 (NS, non-significant). For most behavioural assays we used a two-way ANOVA with a Bonferroni multiple comparison post-hoc test unless otherwise specified in the figure legend or Supplementary Table 4. For quantification of western blots for 4EGI-1 or knockdown experiments, we used a two-way ANOVA with Bonferroni's post-hoc test. To compare changes in group means in amplitude or frequency of mEPSCs or mIPSCs, a Student's *t*-test for independent measures was used. A Kolmogorov–Smirnov test was used to compare cumulative distributions of mEPSC and mIPSC parameters between groups (150 events per neuron) and the relative change in cumulative sum of mEPSC or mIPSC total charge transfer of *Eif4ebp2* knockout (expressed as percentage of mean wild type). For fear conditioning, statistical analyses were done by Student's *t*-tests and two-way ANOVA followed by between-group comparisons using Tukey's post-hoc test. For field potential recordings, we used a two-way ANOVA test.

43. Silverman, J. L., Yang, M., Lord, C. & Crawley, J. N. Behavioural phenotyping assays for mouse models of autism. *Nature Rev. Neurosci.* **11**, 490–502 (2010).
44. Hoeffer, C. A. *et al.* Removal of FKBP12 enhances mTOR-Raptor interactions, LTP, memory, and perseverative/repetitive behavior. *Neuron* **60**, 832–845 (2008).
45. Bidinosti, M. *et al.* Postnatal deamidation of 4E–BP2 in brain enhances its association with raptor and alters kinetics of excitatory synaptic transmission. *Mol. Cell* **37**, 797–808 (2010).
46. Petroulakis, E. *et al.* p53-dependent translational control of senescence and transformation via 4E-BPs. *Cancer Cell* **16**, 439–446 (2009).

LETTER

Mammalian heart renewal by pre-existing cardiomyocytes

Samuel E. Senyo[1], Matthew L. Steinhauser[1], Christie L. Pizzimenti[1], Vicky K. Yang[1], Lei Cai[1], Mei Wang[4,5], Ting-Di Wu[2,3], Jean-Luc Guerquin-Kern[2,3], Claude P. Lechene[4,5] & Richard T. Lee[1,6]

Although recent studies have revealed that heart cells are generated in adult mammals, the frequency of generation and the source of new heart cells are not yet known. Some studies suggest a high rate of stem cell activity with differentiation of progenitors to cardiomyocytes[1]. Other studies suggest that new cardiomyocytes are born at a very low rate[2–4], and that they may be derived from the division of pre-existing cardiomyocytes. Here we show, by combining two different pulse–chase approaches—genetic fate-mapping with stable isotope labelling, and multi-isotope imaging mass spectrometry—that the genesis of cardiomyocytes occurs at a low rate by the division of pre-existing cardiomyocytes during normal ageing, a process that increases adjacent to areas of myocardial injury. We found that cell cycle activity during normal ageing and after injury led to polyploidy and multinucleation, but also to new diploid, mononucleate cardiomyocytes. These data reveal pre-existing cardiomyocytes as the dominant source of cardiomyocyte replacement in normal mammalian myocardial homeostasis as well as after myocardial injury.

Despite intensive research, fundamental aspects of the mammalian heart's capacity for self-renewal are still being actively debated[5,6]. Estimates of cardiomyocyte turnover range from less than 1% per year[2–4] to more than 40% per year[7]. Turnover has been reported to either decrease[3] or increase[7] with age, and the source of new cardiomyocytes has been attributed both to the division of existing myocytes[8] and to progenitors residing within the heart[9] or in exogenous niches such as bone marrow[10]. Controversy persists about the plasticity of the adult heart; in part this controversy is due to methodological challenges associated with studying slowly replenished tissues. Toxicity attributed to radiolabelled thymidine[11] and halogenated nucleotide analogues[12] limits the duration of labelling and may produce direct biological effects. Tissue autofluorescence can decrease the sensitivity and specificity of immunofluorescent methods of detecting cell cycle activity[5,13], such as probing for transient expression of cell cycle markers or the incorporation of halogenated nucleotide analogues. The challenge of measuring cardiomyocyte turnover is further compounded by the faster rate of turnover of cardiac stromal cells relative to that of cardiomyocytes[14].

Here we used multi-isotope imaging mass spectrometry (MIMS) to study cardiomyocyte turnover and to determine whether new cardiomyocytes are derived from pre-existing myocytes or from a progenitor pool (Fig. 1a). MIMS uses ion microscopy and mass spectrometry to generate high-resolution quantitative mass images and localize stable isotope reporters in domains smaller than $1\,\mu m^3$ (refs 15–17). MIMS generates ^{14}N quantitative mass images by measuring the atomic composition of the sample surface with a lateral resolution of less than 50 nm and a depth resolution of a few atomic layers. Cardiomyocyte cell borders and intracellular organelles were easily resolved (Fig. 1b). Regions of interest could be analysed at higher resolution, revealing cardiomyocyte-specific subcellular ultrastructure, including

sarcomeres (Fig. 1c and Supplementary Fig. 1a). In all subsequent analyses, cardiomyocyte nuclei were identified by their location within sarcomere-containing cells, distinguishing them from adjacent stromal cells.

A great advantage of MIMS is the detection of nonradioactive stable isotope tracers. As an integral part of animate and inanimate matter, they do not alter biochemical reactions and are not harmful to the organism[18]. MIMS localizes stable isotope tracers by simultaneously quantifying multiple masses from each analysed domain; this enables the generation of a quantitative ratio image of two stable isotopes of the same element[15]. The incorporation of a tracer tagged with the rare stable isotope of nitrogen (^{15}N) is detectable with high precision by an increase in $^{15}N{:}^{14}N$ above the natural ratio (0.37%). Nuclear incorporation of [^{15}N]thymidine is evident in cells that have divided during a 1-week labelling period, as observed in the small-intestinal epithelium, which turns over completely in 3–5 days[16] (Fig. 1d); in contrast, [^{15}N]thymidine-labelled cells are rarely observed in the heart (Fig. 1e) after 1 week of labelling. In subsequent studies, small intestine was used as a positive control to confirm label delivery.
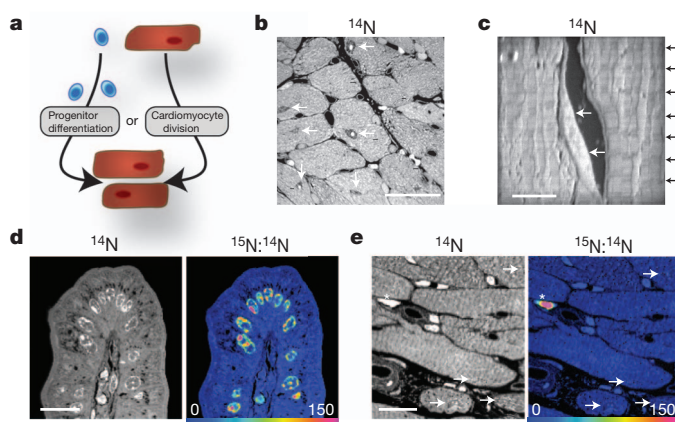

Figure 1 | Use of MIMS to study cardiomyocyte turnover. a, Primary question: are new cardiomyocytes derived from progenitors or from pre-existing cardiomyocytes? **b**, ^{14}N mass image. Subcellular details are evident, including cardiomyocyte nuclei (white arrows). Scale bar, 20 µm. **c**, MIMS resolves periodic sarcomeres (black arrows) in cardiomyocytes. Non-cardiomyocytes (white arrows) are seen outside cardiomyocyte borders. Scale bar, 5 µm. **d**, Right, $^{15}N{:}^{14}N$ hue–saturation–intensity image of small-intestinal epithelium after labelling with [^{15}N]thymidine. The scale ranges from blue, where the ratio is equivalent to natural ratio (0.37%, expressed as 0% above natural ratio (enrichment over natural ratio)), to red, where the ratio is 150% above natural ratio. ^{15}N labelling is concentrated in nuclei in a pattern resembling chromatin. Scale bar, 15 µm. **e**, Right, $^{15}N{:}^{14}N$ hue–saturation–intensity image of heart section (left ventricle). [^{15}N]Thymidine was administered for 1 week. Asterisk, rare $^{15}N^+$ interstitial cells. Cardiomyocyte nuclei (white arrows) are unlabelled. Scale bar, 15 µm.

[1]Cardiovascular Division, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Cambridge, Massachusetts 02139, USA. [2]INSERM U.759, 91405 Orsay, France. [3]Institut Curie, Laboratoire de Microscopie Ionique, 91405 Orsay, France. [4]National Resource for Imaging Mass Spectrometry, Cambridge, Massachusetts 02139, USA. [5]Genetics Division, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Cambridge, Massachusetts 02139, USA. [6]Harvard Stem Cell Institute, Cambridge, Massachusetts 02139, USA.
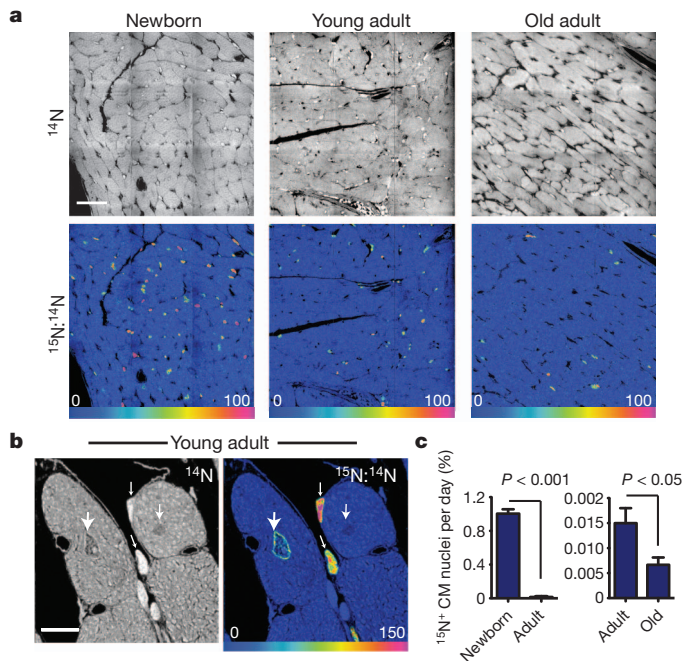
Figure 2 | Cardiomyocyte DNA synthesis decreases with age.
a, [15N]Thymidine was administered for 8 weeks to mice of different ages: newborn, starting at postnatal day 4; young adult, starting at 2 months; old adult, starting at 22 months. Top, 14N mass images show histological details. Bottom, 15N:14N hue–saturation–intensity images show 15N+ nuclei. Mosaics are constructed from nine tiles, 60 μm each. Scale bar, 30 μm. **b**, High-magnification analysis shows a cardiomyocyte from the young adult with nuclear 15N labelling (large arrow), two labelled non-cardiomyocytes (small arrows) and an adjacent unlabelled cardiomyocyte nucleus (medium arrow). Scale bar, 10 μm. **c**, Age-related decline in cardiomyocyte DNA synthesis. Left, comparison of newborn with young adult. Right, scale reduced to compare young adult with old adult ($n = 3$ mice per group). Error bars indicate s.e.m.

To evaluate an age-related change in cell cycle activity, we administered [15N]thymidine for 8 weeks to three age groups of C57BL6 mice starting at day 4 (newborn), at 10 weeks (young adult) and at 22 months (old adult) (Supplementary Fig. 2). We then performed MIMS analysis (Fig. 2a, b and Supplementary Fig. 3). In the newborn group, 56% ± 3% (mean ± s.e.m.; $n = 3$ mice) of cardiomyocytes showed 15N nuclear labelling, consistent with the well-accepted occurrence of cardiomyocyte DNA synthesis during postnatal development[19]. We observed a marked decline in the frequency of 15N-labelled cardiomyocyte nuclei (15N+ CM) in the young adult (newborn, 1.00% ± 0.05% 15N+ CM per day; young adult, 0.015% ± 0.003% 15N+ CM per day (means ± s.e.m.); $n = 3$ mice per group, $P < 0.001$) (Fig. 2a, c and Supplementary Fig. 3). We found a further decrease in cardiomyocyte DNA synthesis in old mice (young adult, 0.015% ± 0.003% 15N+ CM per day; old adult, 0.007% ± 0.002% 15N+ CM per day (means ± s.e.m.); $n = 3$ mice per group, $P < 0.05$) (Fig. 2c). The observed pattern of 15N labelling in cardiomyocyte nuclei is consistent with the known chromatin distribution pattern in cardiomyocytes[20] (Supplementary Fig. 1b) and was measured at levels that could not be explained by DNA repair (Supplementary Fig. 4). Extrapolating DNA synthesis measured in cardiomyocytes over 8 weeks yields a yearly rate of 5.5% in the young adults and 2.6% in the old mice. Given that cardiomyocytes are known to undergo DNA replication without completing the cell cycle[19,21,22], these calculations represent the upper limit of cardiomyocyte generation under normal homeostatic conditions, indicating a low rate of cardiogenesis.

To test whether cell cycle activity occurred in pre-existing cardiomyocytes or was dependent on a progenitor pool, we performed [15N]thymidine labelling of double-transgenic MerCreMer/ZEG mice, previously developed for genetic lineage mapping (Fig. 3a)[23,24].
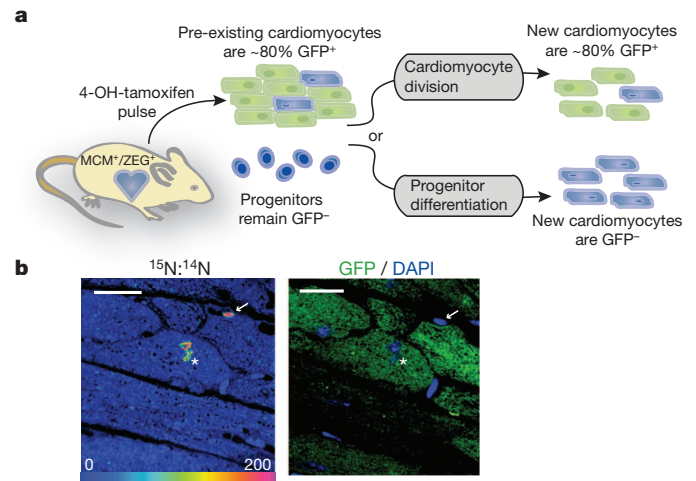
Figure 3 | New cardiomyocytes are derived from pre-existing cardiomyocytes during ageing. a, Experimental strategy. MerCreMer+/ZEG+ (MCM+ZEG+) mice ($n = 4$) were treated for 2 weeks with 4-OH-tamoxifen to induce cardiomyocyte-specific GFP expression. [15N]Thymidine was administered continuously during a 10-week chase, then cycling cells were identified by 15N labelling. New cardiomyocytes (15N+) derived from pre-existing cardiomyocytes should express GFP at a rate similar to that of the surrounding quiescent (15N−) cardiomyocytes. New cardiomyocytes (15N+) derived from progenitors should be GFP−. **b**, Left, 15N:14N hue–saturation–intensity image showing a [15N]thymidine-labelled cardiomyocyte nucleus (white asterisk) and a 15N+ non-cardiomyocyte (white arrow). Right, immunofluorescent image showing that the 15N+ cardiomyocyte is GFP+. Scale bars, 15 μm.

MerCreMer/ZEG cardiomyocytes irreversibly express green fluorescent protein (GFP) after treatment with 4-OH-tamoxifen, allowing pulse labelling of existing cardiomyocytes with a reproducible efficiency of about 80%. Although some have reported rare GFP expression by non-cardiomyocytes with this approach[25], we did not detect GFP expression in interstitial cells isolated from MerCreMer/ZEG hearts, nor did we detect GFP expression by Sca1 or c-Kit-expressing progenitors in histological sections (Supplementary Fig. 5). Thus, during a chase period, cardiomyocytes generated from progenitors should be GFP−, whereas cardiomyocytes arising from pre-existing cardiomyocytes should express GFP at a frequency similar to the background rate induced by 4-OH-tamoxifen. We administered 4-OH-tamoxifen for 2 weeks to 8-week-old mice ($n = 4$); during a subsequent 10-week chase, mice received [15N]thymidine by means of an osmotic minipump (20 μg h−1). With MIMS analysis, we identified 35 15N+ cardiomyocytes (of 4,190 analysed) over 10 weeks, yielding a projected yearly rate of cardiomyocyte DNA replication of 4.4%. Extrapolating from previous reports of high stem-cell-dependent cardiomyocyte turnover[7], we had expected to detect more than 320 cardiomyocytes entering the cell cycle; these results exclude such a high rate of turnover (expected 15N+ CM = 321; observed = 35, Fisher's exact $P < 0.0001$). Immunofluorescent staining for GFP was performed on adjacent sections, and an observer unaware of MIMS analysis results assessed GFP status. Of 15N+ cardiomyocytes, 77% expressed GFP, a frequency essentially identical to that of surrounding 15N− cardiomyocytes (15N+ CM, 77%; 15N− CM, 84%, Fisher's exact P = n.s.) (Table 1). If new cardiomyocytes were derived from progenitors, 15N+ cardiomyocytes would have been GFP− (expected = 0/35; observed = 27/35, Fisher's exact $P < 0.0001$). These data show that 15N+ cardiomyocytes resulted from DNA synthesis by pre-existing cardiomyocytes and exclude a substantial contribution from stem cells in cardiomyocyte replacement in the uninjured heart.

Cardiomyocytes can undergo DNA replication without completing the cell cycle. Although multinucleation and polyploidization occur during early postnatal development and in response to myocardial

Table 1 | $^{15}N^+$ cardiomyocytes ($^{15}N^+$ CM) during normal ageing

	Total counted	GFP$^+$	Percentage GFP$^+$
$^{15}N^-$ CM	1,111	933	84
$^{15}N^+$ CM	35	27	77
Polyploid CM	12	10	83
Diploid CM	22	17	77
Multinucleate CM	18	14	78
Mononucleate CM	17	14	82
Diploid/mononucleate CM	6	5	83

The results include a new diploid/mononucleate fraction, demonstrating a similar rate of GFP expression compared with background $^{15}N^-$ cardiomyocytes ($^{15}N^-$ CM).

stress[2,19], we considered the possibility that these processes could account for $^{15}N^+$ cardiomyocytes in the uninjured adult mouse. We performed fluorescent *in situ* hybridization in adjacent sections to assess the ploidy state of each $^{15}N^+$ cardiomyocyte and surrounding $^{15}N^-$ cardiomyocytes, and an observer unaware of the results of MIMS analysis identified fluorescently labelled chromosomes (Supplementary Fig. 6). Although we found $^{15}N^+$ cardiomyocytes that were polyploid ($4n$ or greater), we observed a higher frequency of diploid nuclei in the $^{15}N^+$ pool than in surrounding $^{15}N^-$ cardiomyocytes ($^{15}N^+$ diploid:polyploid = 22:12; $^{15}N^-$ diploid:polyploid = 9:56, Fisher's exact $P < 0.0001$), consistent with ongoing cell division. We then assessed each cell for multinucleation by using serial 0.5-μm sections adjacent to the section used for MIMS analysis (Supplementary Fig. 6). We observed that 49% of $^{15}N^+$ cardiomyocytes were mononucleate, in contrast with a frequency of 24% for surrounding $^{15}N^-$ cardiomyocytes (Fisher's exact $P < 0.01$), also consistent with cell division. Most cardiomyocyte DNA synthesis occurred in polyploid and/or multinucleate cardiomyocytes, as might be expected with a physiological hypertrophic response, and was thus unlikely to indicate cardiomyocyte division; however, 17% (6 of 35 $^{15}N^+$ CM) were diploid and mononucleate, consistent with newly generated cardiomyocytes (Supplementary Fig. 7). The mononucleate, diploid, $^{15}N^+$ cardiomyocytes

were also predominantly GFP$^+$ (5 of 6 = 83%, compared with 82% background frequency, Fisher's exact P = n.s.), suggesting that they arose from pre-existing cardiomyocytes at a slow annual rate of 0.76% (n = 6 of 4,190 over 10 weeks).

We next used MIMS and genetic fate mapping to study myocardial injury. Cardiomyocyte GFP labelling was induced in MerCreMer/ZEG mice with 4-OH-tamoxifen. Mice then underwent experimental myocardial infarction or sham surgery followed by continuous labelling with [^{15}N]thymidine for 8 weeks. The frequency of $^{15}N^+$ cardiomyocytes in sham-operated mice was similar to that in previous experiments in unoperated mice (yearly projected rates: sham, 6.8%; unoperated, 4.4%), but increased significantly adjacent to infarcted myocardium (total $^{15}N^+$ cardiomyocyte nuclei: myocardial infarction, 23.0%; sham, 1.1%; Fig. 4a, b and Supplementary Fig. 8). We examined the nucleation, ploidy status and GFP expression of $^{15}N^+$ cardiomyocytes and surrounding unlabelled cardiomyocytes. We found a significant dilution of the GFP$^+$ cardiomyocyte pool at the border region, as previously shown[23,24] (67% versus 79%, $P < 0.05$; Table 2 and Supplementary Fig. 9); however, $^{15}N^+$ myocytes demonstrated a similar frequency of GFP expression to that in unlabelled myocytes (71% versus 67%, Fisher's exact P = n.s.), suggesting that DNA synthesis was primarily occurring in pre-existing cardiomyocytes. Of $^{15}N^+$ cardiomyocytes, about 14% were mononucleate and diploid, consistent with division of pre-existing cardiomyocytes (Supplementary Figs 6 and 7). We observed a higher DNA content (more than $2n$) in the remaining cardiomyocytes, as expected with compensatory hypertrophy after injury. Thus, in the 8 weeks after myocardial infarction, roughly 3.2% of the cardiomyocytes adjacent to the infarct had unambiguously undergone division (total $^{15}N^+ \times$ mononucleate diploid fraction = 23% × 0.14 = 3.2%). The low rate of cardiomyocyte cell cycle completion was further supported by the absence of detectable aurora B kinase, a transiently expressed cytokinesis marker, which was detected in rapidly proliferating small intestinal cells but not in cardiomyocytes
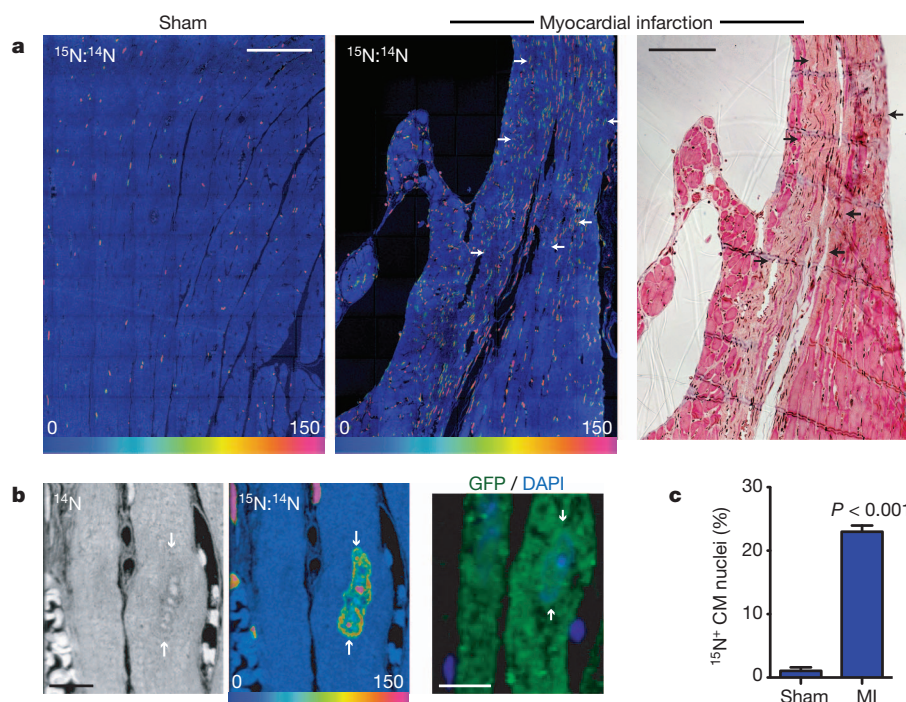

Figure 4 | Myocardial injury stimulates division of pre-existing cardiomyocytes. a, Myocardial infarction leads to extensive DNA synthesis within and adjacent to a scar (arrows). MerCreMer$^+$/ZEG$^+$ mice were treated for 2 weeks with 4-OH-tamoxifen to induce cardiomyocyte-specific GFP expression before myocardial infarction or sham surgery, then [^{15}N]thymidine was administered continuously for 8 weeks. Mosaics of 70 MIMS tiles (each 60 μm × 60 μm) are shown. Trichrome-stained adjacent section (right) shows the scar (black arrows). Scale bars, 90 μm. **b**, [^{15}N]Thymidine-labelled cardiomyocyte nucleus (white arrows) from myocardial infarction border region. Immunofluorescent staining shows that the cardiomyocyte is GFP$^+$. Scale bars, 10 μm. **c**, Mean percentage of $^{15}N^+$ cardiomyocyte nuclei after myocardial infarction (MI; n = 4) in the scar border region compared with that in sham-operated mice (n = 3). Error bars indicate s.e.m.

Table 2 | GFP analysis of ^{15}N$^+$ cardiomyocytes (^{15}N$^+$ CM)

	Total counted	GFP$^+$	Percentage GFP$^+$
Sham	3,411	2,693	79
^{15}N$^+$ CM	23	19	83
Myocardial infarction	7,063	4,766	67
^{15}N$^+$ CM	205	146	71
Polyploid CM	126	97	77
Diploid CM	65	47	72
Multinucleate CM	59	42	71
Mononucleate CM	41	29	71
Diploid/mononucleate CM	16	11	69

Pooled analysis of four mice in the myocardial infarction group and three mice in the sham group. Myocardial infarction led to dilution of GFP$^+$ cardiomyocytes (CM) adjacent to the scar. New cardiomyocytes generated after myocardial infarction (the ^{15}N$^+$/diploid/mononucleate pool) were predominantly GFP$^+$, which is consistent with the division of pre-existing cardiomyocytes.

(Supplementary Fig. 10). We also considered the possibility that a subset of ^{15}N$^+$ myocytes that were multinucleate and/or polyploid resulted from division followed by additional rounds of DNA synthesis without division. However, quantitative analysis of the ^{15}N$^+$ population did not identify a subpopulation that had accumulated additional ^{15}N-label, in contrast with what would be expected in such circumstances (Supplementary Fig. 11). Taken together, these data suggest that adult cardiomyocytes retain some capacity to re-enter the cell cycle, but that most DNA synthesis after injury occurs in pre-existing cardiomyocytes without completion of cell division.

If dilution of the GFP$^+$ cardiomyocyte pool cannot be attributed to division and differentiation of endogenous progenitors, do these data exclude a role for progenitors in the adult mammalian heart? The data could be explained by preferential loss of GFP$^+$ cardiomyocytes after injury, a process that we have previously considered but for which we have not found supporting evidence[23]. Such an explanation excludes a role for endogenous progenitors in cardiac repair and would be consistent with data emerging from lower vertebrates[8,26] and the neonatal mouse[27] in which pre-existing cardiomyocytes are the cellular source for cardiomyocyte repletion. A second possibility to explain the dilution of the GFP$^+$ cardiomyocyte pool is that injury stimulates progenitor differentiation without division; inevitably, this would lead to exhaustion of the progenitor pool, which if true could explain the limited regenerative potential of the adult mammalian heart.

This study demonstrates the birth of cardiomyocytes from pre-existing cardiomyocytes at a projected rate of roughly 0.76% per year (^{15}N$^+$ annual rate × mononucleate diploid fraction = 4.4% × 0.17) in the young adult mouse under normal homeostatic conditions, a rate that declines with age but increases after myocardial injury in the border region. This study shows that cardiac progenitors do not have a significant function in myocardial homeostasis in mammals and suggests that their role after injury is also limited.

METHODS SUMMARY

Multi-isotope imaging mass spectrometry (MIMS) was performed as described previously[16]. Mice were used in accordance with the Guide for the Use and Care of Laboratory Animals and experiments were approved by the Harvard Medical School Standing Committee on Animals. Male C57Bl/6 mice were used for the ageing experiment (Charles River). Double transgenic male MerCreMer/ZEG mice[23] were used in all other experiments.

Full Methods and any associated references are available in the online version of the paper.

1. Hosoda, T. *et al.* Clonality of mouse and human cardiomyogenesis *in vivo. Proc. Natl Acad. Sci. USA* **106,** 17169–17174 (2009).
2. Soonpaa, M. H. & Field, L. J. Assessment of cardiomyocyte DNA synthesis in normal and injured adult mouse hearts. *Am. J. Physiol. Heart Circ. Physiol.* **272,** H220–H226 (1997).
3. Bergmann, O. *et al.* Evidence for cardiomyocyte renewal in humans. *Science* **324,** 98–102 (2009).
4. Walsh, S., Ponten, A., Fleischmann, B. K. & Jovinge, S. Cardiomyocyte cell cycle control and growth estimation *in vivo*—an analysis based on cardiomyocyte nuclei. *Cardiovasc. Res.* **86,** 365–373 (2010).
5. Laflamme, M. A. & Murry, C. E. Heart regeneration. *Nature* **473,** 326–335 (2011).
6. Hosoda, T., Rota, M., Kajstura, J., Leri, A. & Anversa, P. Role of stem cells in cardiovascular biology. *J. Thromb. Haemost.* **9** (Suppl. 1), 151–161 (2011).
7. Kajstura, J. *et al.* Myocyte turnover in the aging human heart. *Circ. Res.* **107,** 1374–1386 (2010).
8. Kikuchi, K. *et al.* Primary contribution to zebrafish heart regeneration by *gata4$^+$* cardiomyocytes. *Nature* **464,** 601–605 (2010).
9. Beltrami, A. P. *et al.* Adult cardiac stem cells are multipotent and support myocardial regeneration. *Cell* **114,** 763–776 (2003).
10. Orlic, D. *et al.* Bone marrow cells regenerate infarcted myocardium. *Nature* **410,** 701–705 (2001).
11. Hu, V. W., Black, G. E., Torres-Duarte, A. & Abramson, F. P. ^{3}H-thymidine is a defective tool with which to measure rates of DNA synthesis. *FASEB J.* **16,** 1456–1457 (2002).
12. Wilson, A. *et al.* Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* **135,** 1118–1129 (2008).
13. Laflamme, M. A. & Murry, C. E. Regenerating the heart. *Nature Biotechnol.* **23,** 845–856 (2005).
14. Bergmann, O. *et al.* Identification of cardiomyocyte nuclei and assessment of ploidy for the analysis of cell turnover. *Exp. Cell Res.* **317,** 188–194 (2010).
15. Lechene, C. *et al.* High-resolution quantitative imaging of mammalian and bacterial cells using stable isotope mass spectrometry. *J. Biol.* **5,** 1–30 (2006).
16. Steinhauser, M. L. *et al.* Multi-isotope imaging mass spectrometry quantifies stem cell division and metabolism. *Nature* **481,** 516–519 (2012).
17. Zhang, D. S. *et al.* Multi-isotope imaging mass spectrometry reveals slow protein turnover in hair-cell stereocilia. *Nature* **481,** 520–524 (2012).
18. Klein, P. D. & Klein, E. R. Stable isotopes: origins and safety. *J. Clin. Pharmacol.* **26,** 378–382 (1986).
19. Li, F., Wang, X., Capasso, J. M. & Gerdes, A. M. Rapid transition of cardiac myocytes from hyperplasia to hypertrophy during postnatal development. *J. Mol. Cell. Cardiol.* **28,** 1737–1746 (1996).
20. Nikolova, V. *et al.* Defects in nuclear structure and function promote dilated cardiomyopathy in lamin A/C-deficient mice. *J. Clin. Invest.* **113,** 357–369 (2004).
21. Adler, C. P. & Friedburg, H. Myocardial DNA content, ploidy level and cell number in geriatric hearts: post-mortem examinations of human myocardium in old age. *J. Mol. Cell. Cardiol.* **18,** 39–53 (1986).
22. Katzberg, A. A., Farmer, B. B. & Harris, R. A. The predominance of binucleation in isolated rat heart myocytes. *Am. J. Anat.* **149,** 489–499 (1977).
23. Hsieh, P. C. *et al.* Evidence from a genetic fate-mapping study that stem cells refresh adult mammalian cardiomyocytes after injury. *Nature Med.* **13,** 970–974 (2007).
24. Loffredo, F. S., Steinhauser, M. L., Gannon, J. & Lee, R. T. Bone marrow-derived cell therapy stimulates endogenous cardiomyocyte progenitors and promotes cardiac repair. *Cell Stem Cell* **8,** 389–398 (2011).
25. Dong, F. *et al.* Myocardial CXCR4 expression is required for mesenchymal stem cell mediated repair following acute myocardial infarction. *Circulation* **126,** 314–324 (2012).
26. Jopling, C. *et al.* Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature* **464,** 606–609 (2010).
27. Porrello, E. R. *et al.* Transient regenerative potential of the neonatal mouse heart. *Science* **331,** 1078–1080 (2011).

Author Contributions S.E.S., M.L.S. and R.T.L. designed the experiments. S.E.S., M.L.S. and V.K.Y. performed the experiments. L.C., C.L.P., V.K.Y., M.L.S. and S.E.S. performed histology. M.W. and S.E.S. operated the prototype nanoSIMS instrument. T.-D.W. operated the nanoSIMS instrument at Institut Curie with input from J.-L.G.K; S.E.S. and M.L.S. analysed the data, with input from C.P.L. and R.T.L. C.P.L. supervised all MIMS analyses and provided critical feedback at all junctures. S.E.S. and M.L.S. made the figures. S.E.S., M.L.S. and R.T.L. wrote the manuscript. All authors approved the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.T.L. (rlee@partners.org).

METHODS

Mice. All experiments were conducted in accordance with the Guide for the Use and Care of Laboratory Animals and approved by the Harvard Medical School Standing Committee on Animals. C57Bl/6 male mice were obtained from Charles River.

We generated double transgenic MerCreMer/ZEG male mice by crossbreeding cardiomyocyte-specific MerCreMer mice and ZEG mice (Jackson Laboratory). β-Galactosidase–GFP is under the control of a cytomegalovirus (CMV) enhancer/chicken actin promoter (*Actb*); the background strain was C57BL/6J (N7). The background strain of the MerCreMer mice was C57Bl/6SV129. Genotyping was performed by PCR on tail DNA using the following primers: MerCreMer, 5′-GTCTGACTAGGTGTCCTTCT-3′ (forward) and 5′-CGTCCTCCTGCTGGTA TAG-3′ (backward); ZEG, 5′-AAGTTCATCTGCACCACCG-3′ (forward) and 5′-TCCTTGAAGAAGATGGTGCG-3′ (backward); ZEG control, 5′-CTAGGC CACAGAATTGAAAGATCT-3′ (forward) and 5′-GTAGGTGGAAATTCTAG CATCATCC-3′ (backward). To induce Cre recombination and GFP labelling in cardiomyocytes, we injected 4-OH-tamoxifen (a gift from Laboratoires Besins), dissolved in peanut oil (Sigma) at a concentration of 5 mg ml^{-1}, intraperitoneally into 8-week-old MerCreMer/ZEG mice daily at a dosage of 0.5 mg per day for 14 days.

Experimental myocardial infarction. Mice were subjected to experimental myocardial infarction as described. Surgeries were performed by a single operator with more than 20 years of experience in the performance of coronary ligation in rodents. In brief, the left coronary artery was permanently ligated roughly 2 mm below the left atrial appendage. For sham operations, the thoracic cavity was opened and the heart exposed, but no intramyocardial sutures were placed. [^{15}N]Thymidine (Cambridge Isotopes) was administered at a rate of 20 μg h^{-1} with an osmotic minipump (Alzet), implanted subcutaneously at the time of experimental myocardial infarction after a single intraperitoneal bolus dose of 500 μg.

MIMS data acquisition. The factory prototype of the NanoSIMS50 as well as a standard NanoSIMS 50 and a large-radius NanoSIMS 50L (Cameca) was used for MIMS analysis as described previously[15]. A focused beam of Cs$^+$ ions was used to sputter a few atomic layers and generate secondary ions from the left ventricular free wall of heart section samples. The Cs$^+$ primary ions were scanned over a raster pattern of either 256 pixels × 256 pixels or 512 pixels × 512 pixels. At each pixel location, the secondary-ion intensities for ^{12}C$^-$, ^{13}C$^-$, ^{12}C^{14}N$^-$ and ^{12}C^{15}N$^-$ were recorded in parallel from the same sputtered volume. The detection of nitrogen requires the use of cluster ions ^{12}C^{14}N$^-$ and ^{12}C^{15}N$^-$ for ^{14}N and ^{15}N, respectively, as a result of the low ionization efficiency of nitrogen as N$^-$.

MIMS data analysis. From a single field image acquisition, we first extracted four image files: the four original quantitative mass images (QMIs; ^{12}C, ^{13}C, ^{12}C^{14}N and ^{12}C^{15}N) and the two ratio images (^{13}C:^{12}C and ^{12}C^{15}N:^{12}C^{14}N), derived from the pixel-by-pixel division of the ^{13}C QMI by the ^{12}C QMI and of the ^{12}C^{15}N QMI by the ^{12}C^{14}N QMI, respectively. We then used a hue–saturation–intensity transformation of the ratio image to map ^{15}N-labelled regions. The hue corresponds to the ratio value, and the intensity at a given hue is an index of statistical reliability.

[^{15}N]Thymidine labelling. For the neonatal cohort in the ageing experiment, labelling was started at postnatal day 4 with subcutaneous injections of 50 μg g^{-1} [^{15}N]thymidine (Cambridge Isotopes) every 12 h and continued until postnatal week 4. Starting at age 4 weeks—and in all other experiments using adult mice—osmotic minipumps (Alzet) were implanted subcutaneously, delivering [^{15}N]thymidine at a rate of 20 μg h^{-1}.

Multinucleation analysis. Serial adjacent sections (0.5 μm) were stained to identify cardiomyocyte borders. A given cardiomyocyte was tracked in the vertical axis by locating it in serial sections. Uninjured hearts were stained by using a modified PAS protocol with standard solutions (Electron Microscopy Services), but with longer incubation times optimized for LR white embedding. Slides were incubated for 1 h in xylene at 37 °C, rehydrated through graded alcohols, incubated in periodic acid for 2 h and then in Schiff's reagent for two nights. Sections were counterstained with haematoxylin and with Scott's Bluing for 1 h each. Injured hearts were stained with a modified Trichrome staining protocol with standard solutions (Fisher Scientific), but with longer incubation times. Slides were incubated for 1 h in xylene at 37 °C, rehydrated through graded alcohols, incubated for 1 h in Bouin's fluid at 56 °C, rinsed in tap water, incubated for 1 h in Weigert's iron haematoxylin stain, rinsed in tap water, incubated for 1 h in scarlet–acid fuchsin solution, rinsed in distilled deionized water, incubated for 30 min in phophotungistic–phophomolybdic acid solution, incubated for 30 min in aniline blue stain solution, and incubated for 20 min in 1% acetic acid.

Fluorescent *in situ* hybridization. Sections were incubated for 15 min in proteinase K (50 μg ml^{-1}) at 60 °C. After a wash with PBS containing 45 mM MgCl$_2$, slides were postfixed in 4% paraformaldehyde (in PBS/MgCl$_2$) and dehydrated through graded alcohols. Biotinylated-labelled chromosome Y paint (Star-FISH; Cambio) in hybridization mix was applied to sections and sealed under glass with rubber cement (some samples were analysed with chromosome 18 paint because of product discontinuation of chromosome Y paint). Samples were heated to 90 °C for 15 min. After incubation overnight at 37 °C, slides were washed three times with 50% formamide/2 × standard saline citrate at 45 °C, three times with 2 × standard saline citrate at room temperature (25 °C), and twice with 4 × standard saline citrate/0.1% Tween at room temperature. Samples were blocked for 10 min with 4 × standard saline citrate/0.1% Tween/0.05% milk and incubated for 2 h in streptavidin-conjugated Alexa Fluor 488 (Invitrogen) before being washed and mounted. An observer unaware of the MIMS images or [^{15}N]thymidine labelling status of the nuclei assigned ploidy status.

Immunofluorescent staining. Sections were incubated for 5 min in glycine/Tris (50 mM glycine/0.05 M Tris) at room temperature. After a brief wash with 0.05 M Tris, sections were incubated overnight with both chicken anti-GFP (Abcam) and rabbit anti-c-Kit (Abcam) or rat anti-Sca1 (Abcam) with fresh 0.1% BSA in TBS/ 0.1% Tween at 4 °C. After a brief wash with TBS, sections were incubated with anti-chicken Alexa Fluor 488 (Invitrogen) before being washed with TBS and then mounted. An observer unaware of the MIMS images or [^{15}N]thymidine labelling status of the nuclei assigned GFP status.

Fluorescence image analysis. We used a custom-written script in IP Lab version 4.0 (Scanalytics) imaging software for serial image acquisition. Tissue sections were auto-imaged with an Olympus IX-70 microscope with a digital charge-coupled-device camera (CoolSNAP EZ; Roper Scientific), using an automated stage with a piezoelectric objective positioner (Polytec PI; Auburn MA) to compensate for deviations in the z axis. Images were compressed and stitched into a mosaic with stitching software (Canon Photostitch). Multichannel images were merged in ImageJ before stitching.

Statistical analysis. Statistical testing was performed with Prism 3.0 (Graphpad). Results are presented as means ± s.e.m. and were compared by using t-tests (significance was assigned for $P < 0.05$). Data comparing event rates were tested with a Fisher exact test.

Structure and function of the initially transcribing RNA polymerase II–TFIIB complex

Sarah Sainsbury[1], Jürgen Niesser[1] & Patrick Cramer[1]

The general transcription factor (TF) IIB is required for RNA polymerase (Pol) II initiation and extends with its B-reader element into the Pol II active centre cleft. Low-resolution structures of the Pol II–TFIIB complex[1,2] indicated how TFIIB functions in DNA recruitment, but they lacked nucleic acids and half of the B-reader, leaving other TFIIB functions[3,4] enigmatic. Here we report crystal structures of the Pol II–TFIIB complex from the yeast *Saccharomyces cerevisiae* at 3.4 Å resolution and of an initially transcribing complex that additionally contains the DNA template and a 6-nucleotide RNA product. The structures reveal the entire B-reader and protein–nucleic acid interactions, and together with functional data lead to a more complete understanding of transcription initiation. TFIIB partially closes the polymerase cleft to position DNA and assist in its opening. The B-reader does not reach the active site but binds the DNA template strand upstream to assist in the recognition of the initiator sequence and in positioning the transcription start site. TFIIB rearranges active-site residues, induces binding of the catalytic metal ion B, and stimulates initial RNA synthesis allosterically. TFIIB then prevents the emerging DNA–RNA hybrid duplex from tilting, which would impair RNA synthesis. When the RNA grows beyond 6 nucleotides, it is separated from DNA and is directed to its exit tunnel by the B-reader loop. Once the RNA grows to 12–13 nucleotides, it clashes with TFIIB, triggering TFIIB displacement and elongation complex formation. Similar mechanisms may underlie all cellular transcription because all eukaryotic and archaeal RNA polymerases use TFIIB-like factors[5], and the bacterial initiation factor sigma has TFIIB-like topology[1,2] and contains the loop region 3.2 that resembles the B-reader loop in location, charge and function[6–8]. TFIIB and its counterparts may thus account for the two fundamental properties that distinguish RNA from DNA polymerases: primer-independent chain initiation and product separation from the template.

Our previous X-ray analysis of the Pol II–TFIIB complex at 4.3 Å resolution provided a partial backbone model of TFIIB[1]. To obtain a complete and atomic structure, we co-crystallized Pol II with a TFIIB variant lacking the mobile amino-terminal tail and carboxy-terminal cyclin fold (Methods, Supplementary Table and Supplementary Fig. 1). The resulting Pol II–TFIIB structure at 3.4 Å resolution provides details of the interactions of the four TFIIB domains with Pol II: the B-ribbon with the dock, the B-core N-terminal cyclin fold with the wall, the B-reader helix with the RNA exit tunnel, and the B-linker helix with the coiled-coil of the clamp (Fig. 1).

The structure reveals the entire course of the TFIIB polypeptide chain through the Pol II cleft, including the previously lacking[1] B-reader loop (residues 67–79) and a new 'B-reader strand' (residues 80–83). The B-reader loop does not reach the active site, but instead interacts with the Pol II rudder and fork loop 1 to seal part of the cleft. The B-reader strand adds to the two-stranded polymerase lid (Fig. 1a, c), positioning the B-reader and explaining why the lid is essential for transcription[9]. There is no evidence for a B-finger[2].

TFIIB binding induces changes in Pol II domains (Fig. 2). The lobe, protrusion and wall rotate slightly, partially closing the cleft.

The previously mobile[10] tip of the protrusion forms a β-sheet and an ordered loop (α11–α12) that is required for initiation complex stability[11] (Fig. 1d). To induce these changes, the B-ribbon may bind the dock, causing the wall to rotate, which then enables the B-core to bind the wall and protrusion (Figs 1d and 2a). TFIIF also binds the lobe and protrusion[12,13], and stabilizes TFIIB on Pol II[14], in particular binding of the B-core to the wall[15]. These changes contribute to TFIIB function, because mutations in the lobe and protrusion (Supplementary Fig. 2) suppress a TFIIB mutation that causes shifts in the transcription start site (TSS)[16].

The structural changes extend to the polymerase active site, where difference electron density revealed two metal ions with similar occupancy (Fig. 2b). This was consistent with a two-metal ion catalytic mechanism[17,18], but was unexpected because free Pol II only binds metal A, whereas metal B was observed previously only when a nucleoside triphosphate (NTP) was present[19,20]. The two metal ions are bridged by a rotated aspartate D481 side chain (Fig. 2b), which is also observed in structures of NTP-containing elongation complexes of Pol II[20] and bacterial polymerase[21]. The rearrangement abolishes a contact between metal A and residue D483, and re-orientates the backbone carbonyl in the neighbouring Rpb2 residue D837 (Fig. 2b).

This indicates that TFIIB induces active-site rearrangements that increase the affinity for metal B and stimulate catalysis. Consistent with this, the B-reader of the archaeal TFIIB homologue stimulates transcription[22]; truncation of the B-reader loop in human TFIIB impairs extension of a RNA dinucleotide[23]; and the presence of TFIIB and TBP stimulates transcription initiation from pre-opened DNA[15,24]. We find here that TFIIB alone stimulates *de novo* RNA synthesis by Pol II at low or near-physiological NTP concentrations (Fig. 2c and Supplementary Fig. 3). Thus, TFIIB stimulates RNA chain initiation without contacting the catalytic site, by activating Pol II allosterically.

We also solved the structure of an initially transcribing Pol II–TFIIB complex (ITC) by soaking into the crystals a *HIS4* promoter-based scaffold with a downstream DNA duplex and a 6-base-pair (bp) DNA–RNA hybrid. Electron density was observed for the entire RNA, 7 bp of downstream DNA, and the DNA template strand until position −8 upstream of the NTP site +1. The structure at 3.6 Å resolution (Fig. 3 and Supplementary Table) shows again two metal ions bridged by a rotated D481 side chain. Metal A binds the RNA 3′ end and could bind the NTP α-phosphate[21] (Fig. 3b and Supplementary Fig. 4). Metal B is 5 Å away from metal A and may move closer for catalysis upon binding NTP β- and γ-phosphates[21] (Fig. 3b).

The ITC structure indicates that TFIIB stimulates initial RNA synthesis not only by allosteric active-site rearrangement but also by preventing tilting of short DNA–RNA hybrids. Without TFIIB, short hybrids are tilted, and this impairs synthesis, because it occludes the templating DNA base +1 from the active site[25]. Such tilting is prevented in the ITC because TFIIB binds the upstream template strand. B-reader residues R64 and D69 contact template strand bases at positions −7/−8 (Fig. 3c). These interactions may also stabilize the initial DNA bubble during open complex formation.

[1]Gene Center and Department of Biochemistry, Center for Integrated Protein Science CIPSM, Ludwig-Maximilians-Universität München, Feodor-Lynen-Str. 25, 81377 Munich, Germany.
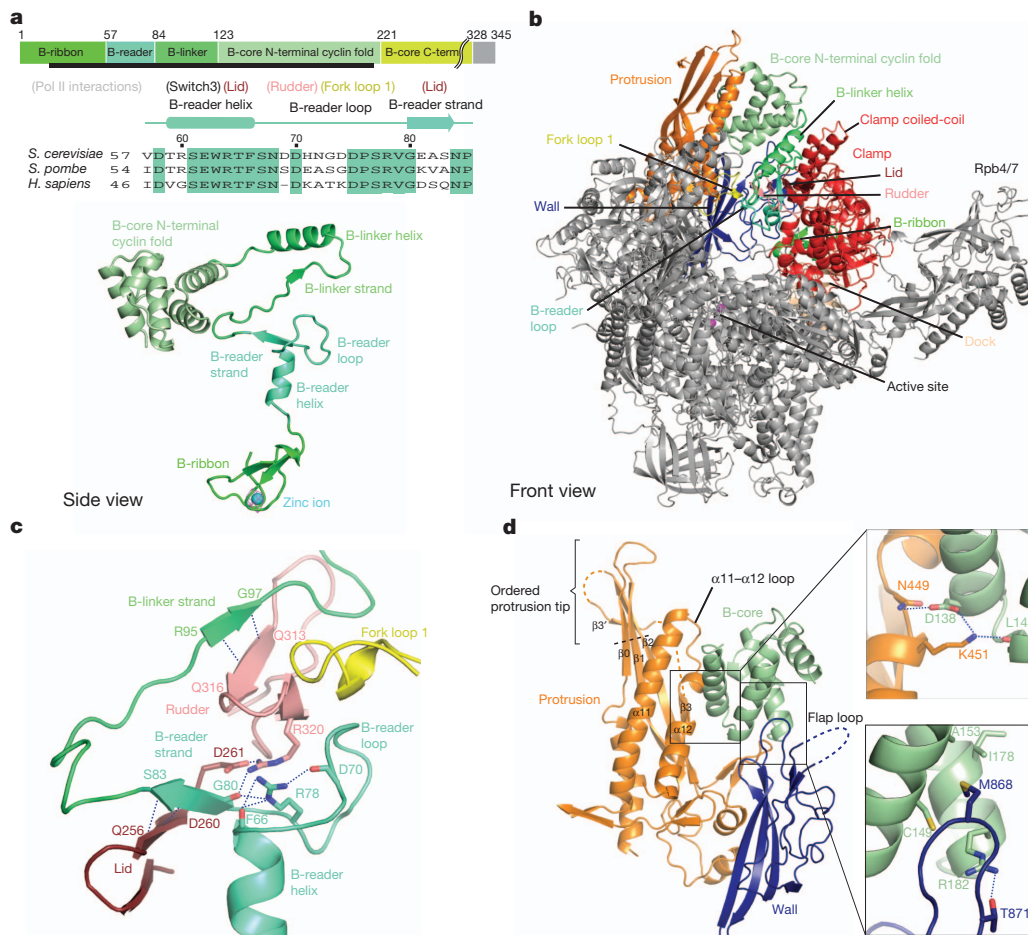
Figure 1 | Structure of the Pol II–TFIIB complex at 3.4 Å resolution.
a, TFIIB domain organization and ribbon model. The colour code is used throughout the figures. Invariant residues are highlighted. **b**, View of Pol II–TFIIB structure showing the active centre cleft. **c**, Interactions of B-linker and B-reader with Pol II rudder, lid and fork loop 1. **d**, Interactions of B-core N-terminal cyclin fold with Pol II protrusion and wall. Previous secondary structure nomenclature was maintained[19] with newly observed elements termed β0 (Rpb2 residues 69–74), β3′ (148–151) and loop α11–α12 (438–443).

The same contacts contribute to the recognition of an initiator (Inr) DNA sequence that defines the TSS and requires a thymine at template position −7/−8 (refs 1, 26, 27). Interaction of an Inr thymine with the B-reader residue R64 (Fig. 3c) apparently occurs in the open DNA complex, because mutation of R64 causes TSS shifts that are sensitive to changes in Inr sequence[27,28]. The B-reader function in Inr recognition is supported by mutational analysis of the invariant TFIIB residue R78, which buttresses the B-reader loop, including the DNA-binding residue D69 (Fig. 1c). Mutations of R78 shift the TSS in yeast[16] and human[29], and can be lethal (Supplementary Fig. 5). Thus, the B-reader adopts a defined structure to assist in TSS selection.

The path of the RNA is blocked by the B-reader loop (Fig. 3). When the RNA grows from 6 to 7 nucleotides, it would clash with three aspartate residues (D70, D74, and D75; Fig. 3c), of which two are conserved among eukaryotes (Fig. 1a). This is consistent with a destabilization of TFIIB binding when RNA grows beyond 6 nucleotides, but TFIIB can be retained beyond this point and is only released when the RNA length reaches 12–13 nucleotides[14]. Thus, the B-reader loop
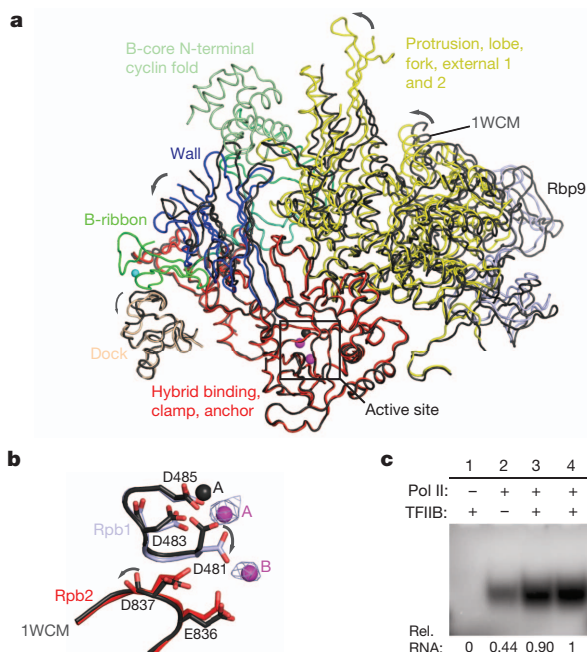
Figure 2 | TFIIB changes Pol II domains and stimulates catalysis. Same colour code as Fig. 1. **a**, Apparent order of TFIIB domain binding and polymerase domain movements. Pol II–TFIIB structure (colours) and free Pol II[10] (dark grey) were superimposed with their Rpb1 subunits. **b**, TFIIB induces active-site rearrangements. Metals A and B are shown as magenta spheres. Difference electron density was contoured at 3.0σ (grey mesh, peak height 4.4σ (metal A) and 3.6σ (metal B)). **c**, TFIIB stimulates de novo RNA synthesis from a HIS4 promoter-derived tailed template. At least twofold stimulation occurs irrespective of whether TFIIB is added before (lane 3) or after (lane 4) DNA.
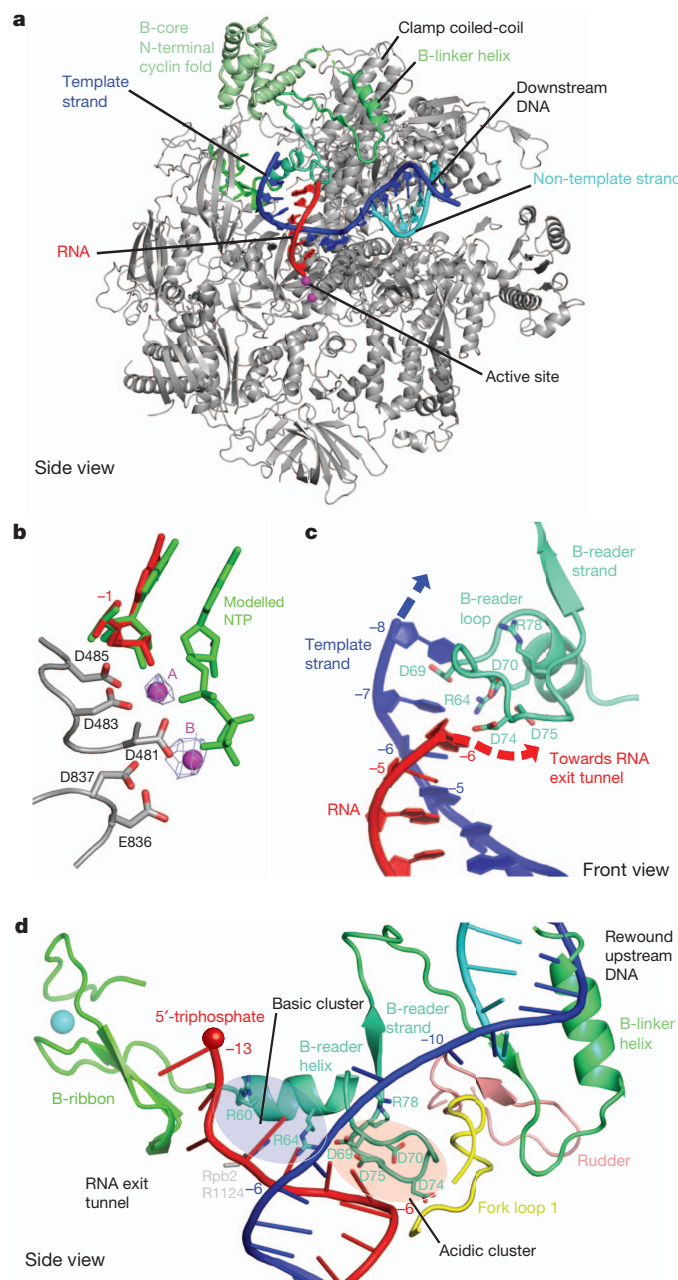
Figure 3 | Structure of initially transcribing Pol II–TFIIB complex (ITC) elucidates RNA initiation and separation, and TFIIB release. Same colour code as Fig. 1. **a**, Ribbon model. The view is related to that in Fig. 1a by a 90° rotation along a vertical axis. **b**, Active-site configuration with RNA 3′-nucleotide in position −1 (red) and superimposed 3′-nucleotide and modelled NTP (green) from a bacterial elongation complex[21]. The metals are covered by difference electron density (grey mesh) contoured at 3.0σ (peak heights 3.8σ (metal A) and 4.5σ (metal B), respectively). **c**, B-reader binds the DNA template strand at position −7/−8 and separates the RNA. **d**, Model of RNA extended into the exit tunnel elucidates TFIIB release.

separates growing RNA from the DNA template and directs it to its exit tunnel. In the elongation complex, RNA separation is achieved by rewinding of the upstream DNA duplex. However, in an ITC with a short RNA that does not yet reach the exit tunnel, such DNA rewinding probably causes collapse of the open DNA bubble, displacing RNA and aborting transcription.

Consistent with a role of the B-reader loop in RNA separation, abortive transcription increases when the loop is destabilized by R78 mutation[16] (Fig. 1c). B-reader loop function is required for transcription because deletion or mutation of B-reader loop residues 69–75

cause defects *in vivo* and *in vitro* (Supplementary Fig. 5), consistent with results for human TFIIB[23] and with *in vitro* data for yeast TFIIB[30]. B-reader loop function is conserved because replacing the loop for its *Schizosaccharomyces pombe* counterpart preserves function (Supplementary Fig. 5).

We collected diffraction data on ITCs with a 6-nucleotide RNA containing an additional 5′-phosphate and with a 7-nucleotide RNA. The data showed that the B-reader loop was unchanged, consistent with its function in RNA separation. Nucleic acids were present, but partial mobility prevented structure completion. The 6-nucleotide RNA 5′-phosphate was revealed beneath the B-reader loop, and guided modelling of an extended RNA.

The model indicates that RNA separation may be achieved by charge repulsion between the RNA 5′-triphosphate and a cluster of acidic residues in the B-reader loop, and by attraction of the triphosphate to a cluster of basic residues in the exit tunnel (Fig. 3d). Alternatively, the RNA 5′-triphosphate may be trapped by B-reader loop aspartates and a bridging metal ion. Our model accommodates 11–12 nucleotides of RNA. Further RNA extension results in a clash with the B-reader, B-ribbon and wall, explaining why TFIIB is released when the RNA length reaches 12–13 nucleotides[14]. TFIIB release enables formation of the mature 8-bp DNA–RNA hybrid observed in the elongation complex.

Bacterial RNA polymerase and its initiation factor sigma apparently use similar mechanisms for transcription initiation. The sigma factor is topologically related to TFIIB[1,2] and contains the loop region 3.2, which resembles the B-reader loop in location and negative charge[6–8]. Region 3.2 is required for formation of the first RNA phosphodiester bond, normal abortive transcription and sigma factor release[6–8].

METHODS SUMMARY

Saccharomyces cerevisiae 12-subunit Pol II (3.0 mg ml^{-1}) was incubated with a fivefold molar excess of TFIIB for 30 min at 298 K before crystallization by vapour diffusion using 32–34% (v/v) tacsimate pH 7.0, 100 mM HEPES pH 7.0, and 5 mM dithiothreitol (DTT) as reservoir solution. Crystals were cryo-protected in mother solution containing 70% (v/v) tacsimate, supplemented with 670 µM *HIS4* tailed template[25] for ITC complexes, and incubated at 281 K over night before flash-cooling in liquid nitrogen. Diffraction data were collected at 100 K at SLS beamline X06SA. Structures were solved by molecular replacement (Protein Data Bank code 1WCM). TFIIB was expressed from a pET21 plasmid in *Escherichia coli* strain Rosetta DE3 (Novagen) and purified by nickel affinity, cation exchange and size-exclusion chromatography. For *in vivo* analysis, TFIIB mutants were cloned into vector pRS315 containing 500 bp upstream and 300 bp downstream of the *SUA7* open reading frame (ORF), and selected on 5-FOA medium[1]. For *in vitro* transcription, nuclear extracts were prepared from wild-type and strain SHY245 (ref. 4) as described and assays performed as reported[1]. For the *de novo* transcription assay, 20 µM rCTP, rGTP and rUTP, and 2 µM rATP supplemented with 5 µCi [α-^{32}P] rATP were incubated with 2.5 pmol Pol II, 25 pmol TFIIB and 5 pmol template for 5 min at 301 K.

Full Methods and any associated references are available in the online version of the paper.

1. Kostrewa, D. *et al.* RNA polymerase II–TFIIB structure and mechanism of transcription initiation. *Nature* **462,** 323–330 (2009).
2. Liu, X., Bushnell, D. A., Wang, D., Calero, G. & Kornberg, R. D. Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. *Science* **327,** 206–209 (2010).
3. Cho, E. J. & Buratowski, S. Evidence that transcription factor IIB is required for a post-assembly step in transcription initiation. *J. Biol. Chem.* **274,** 25807–25813 (1999).
4. Ranish, J. A., Yudkovsky, N. & Hahn, S. Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev.* **13,** 49–63 (1999).
5. Vannini, A. & Cramer, P. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol. Cell* **45,** 439–446 (2012).
6. Vassylyev, D. G. *et al.* Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417,** 712–719 (2002).
7. Murakami, K. S., Masuda, S. & Darst, S. A. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* **296,** 1280–1284 (2002).

8. Kulbachinskiy, A. & Mustaev, A. Region 3.2 of the sigma subunit contributes to the binding of the 3′-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation. *J. Biol. Chem.* **281,** 18273–18276 (2006).

9. Naji, S., Bertero, M. G., Spitalny, P., Cramer, P. & Thomm, M. Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement. *Nucleic Acids Res.* **36,** 676–687 (2008).

10. Armache, K. J., Mitterweger, S., Meinhart, A. & Cramer, P. Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *J. Biol. Chem.* **280,** 7131–7134 (2005).

11. Grünberg, S., Warfield, L. & Hahn, S. Architecture of the RNA polymerase II preinitiation complex and mechanism of ATP-dependent promoter opening. *Nature Struct. Mol. Biol.* **19,** 788–796 (2012).

12. Chen, Z. A. *et al.* Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **29,** 717–726 (2010).

13. Eichner, J., Chen, H. T., Warfield, L. & Hahn, S. Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. *EMBO J.* **29,** 706–716 (2010).

14. Cabart, P., Ujvari, A., Pal, M. & Luse, D. S. Transcription factor TFIIF is not required for initiation by RNA polymerase II, but it is essential to stabilize transcription factor TFIIB in early elongation complexes. *Proc. Natl Acad. Sci. USA* **108,** 15786–15791 (2011).

15. Fishburn, J. & Hahn, S. Architecture of the yeast RNA polymerase II open complex and regulation of activity by TFIIF. *Mol. Cell. Biol.* **32,** 12–25 (2012).

16. Chen, B. S. & Hampsey, M. Functional interaction between TFIIB and the Rpb2 subunit of RNA polymerase II: implications for the mechanism of transcription initiation. *Mol. Cell. Biol.* **24,** 3983–3991 (2004).

17. Steitz, T. A., Smerdon, S. J., Jager, J. & Joyce, C. M. A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science* **266,** 2022–2025 (1994).

18. Sosunov, V. *et al.* Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J.* **22,** 2234–2244 (2003).

19. Cramer, P., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292,** 1863–1876 (2001).

20. Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D. & Kornberg, R. D. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127,** 941–954 (2006).

21. Vassylyev, D. G. *et al.* Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448,** 163–168 (2007).

22. Werner, F. & Weinzierl, R. O. Direct modulation of RNA polymerase core functions by basal transcription factors. *Mol. Cell. Biol.* **25,** 8344–8355 (2005).

23. Tran, K. & Gralla, J. D. Control of the timing of promoter escape and RNA catalysis by the transcription factor IIb fingertip. *J. Biol. Chem.* **283,** 15665–15671 (2008).

24. Pal, M., Ponticelli, A. S. & Luse, D. S. The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. *Mol. Cell* **19,** 101–110 (2005).

25. Cheung, A. C., Sainsbury, S. & Cramer, P. Structural basis of initial RNA polymerase II transcription. *EMBO J.* **30,** 4755–4763 (2011).

26. Zhang, Z. & Dietrich, F. S. Mapping of transcription start sites in *Saccharomyces cerevisiae* using 5′ SAGE. *Nucleic Acids Res.* **33,** 2838–2851 (2005).

27. Kuehner, J. N. & Brow, D. A. Quantitative analysis of *in vivo* initiator selection by yeast RNA polymerase II supports a scanning model. *J. Biol. Chem.* **281,** 14119–14128 (2006).

28. Bangur, C. S., Pardee, T. S. & Ponticelli, A. S. Mutational analysis of the D1/E1 core helices and the conserved N-terminal region of yeast transcription factor IIB (TFIIB): identification of an N-terminal mutant that stabilizes TATA-binding protein-TFIIB-DNA complexes. *Mol. Cell. Biol.* **17,** 6784–6793 (1997).

29. Hawkes, N. A. & Roberts, S. G. The role of human TFIIB in transcription start site selection *in vitro* and *in vivo*. *J. Biol. Chem.* **274,** 14337–14343 (1999).

30. Yang, C. & Ponticelli, A. S. Evidence that RNA polymerase II and not TFIIB is responsible for the difference in transcription initiation patterns between *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Nucleic Acids Res.* **40,** 6495–6507 (2012).

METHODS

Preparation of *S. cerevisiae* Pol II and recombinant TFIIB. *Saccharomyces cerevisiae* 12-subunit Pol II was prepared as described[31]. TFIIB (residues 19–215) was cloned into pET21b with a C-terminal His$_6$ tag and expressed in *E. coli* Rosetta (DE3) (Novagen). The culture was grown in LB medium at 37 °C until absorbance at 600 nm of 0.6, 0.5 mM IPTG was added and the culture was grown for a further 18 h at 20 °C. Cells were collected by centrifugation and frozen in liquid nitrogen. Protein was purified by nickel affinity, cation exchange and size-exclusion chromatography. Cells were lysed by sonication in buffer A (50 mM Tris pH 7.5, 300 mM NaCl, 10 μM ZnCl$_2$, 2 mM dithiothreitol (DTT) supplemented with 5 mM imidazole, 0.2% (v/v) Tween 20 and 1× protease inhibitors (100× stock: 1.42 mg leupeptin, 6.85 mg pepstatin A, 850 mg PMSF, 1,650 mg benzamidine in 50 ml ethanol)). After centrifugation at 30,000*g* for 30 min, the cleared lysate was applied to a pre-equilibrated (buffer A) Ni-NTA agarose column (Qiagen). The column was washed with 10 column volumes of buffer A containing 10 mM imidazole before elution of the protein with buffer A containing 200 mM imidazole. The conductivity of the eluate was adjusted to match that of buffer B (50 mM Tris pH 7.5, 100 mM NaCl, 2 mM DTT) and applied to a MonoS 10/100 GL column (GE healthcare) equilibrated in buffer B. The protein was eluted with a linear gradient from 100 mM to 1 M NaCl. To remove any minor contaminants a final size exclusion step using a Superdex 75 10/300 GL column (GE Healthcare) in 5 mM NaHEPES (pH 7.25 at 20 °C), 40 mM (NH$_4$)$_2$SO$_4$, 10 μM ZnCl$_2$, 10 mM DTT was carried out. The protein was concentrated to 3 mg ml^{-1} in an Amicon Ultra-4 Centrifugal Filter Unit (Millipore) and flash-frozen in liquid nitrogen. TFIIB reader variants were cloned with a C-terminal His$_6$-tag into pOPINE[32] (provided by OPPF-UK) using In-Fusion HD cloning kit (Contech) and purified as described above.

Yeast strains and promoter-dependent *in vitro* transcription assay. Mutations in TFIIB were introduced by PCR. Both wild-type and mutated TFIIB were cloned into pRS315[33] containing 500 bp upstream and 300 bp downstream of the *SUA7* ORF. After transformation into the *SUA7* shuffle strain[1], selection and growth characterization of the respective TFIIB variants was carried out on 5-FOA medium as described[1]. All growth experiments were made as triplicates. Nuclear extracts were prepared from wild-type and strain SHY245 (ref. 4) as reported previously[34]. *In vitro* transcription was essentially performed as described[34], with minor modifications. As template, 150 ng of a plasmid encoding for a *HIS4-SNR14* (pMSe58) fusion promoter was used and the transcription buffer contained 20 mM HEPES pH 7.6, 100 mM potassium acetate, 1 mM EDTA, 5 mM magnesium acetate and 2.5 mM DTT. For activated transcription, 150 ng of the recombinant Gal4–VP16 fusion protein and 10 pmol of the respective recombinant TFIIB variant were used. Primer extension was carried out using 0.125 pmol of a fluorescently labelled DNA primer (5′-Cy5-TTCACCAGTGAGACGGGC AAC). The transcripts were separated using a 6% polyacrylamide/7 M urea TBE gel, scanned with a Typhoon 9410 and quantified with the ImageQuant software (GE healthcare).

Promoter-independent *de novo* transcription assay with 3′-tailed DNA templates. 10 μl reactions of 2.5 pmol Pol II, 25 pmol TFIIB and 5 pmol *HIS4* 3′ tailed-template in 20 mM HEPES pH 7.6, 60 mM (NH$_4$)$_2$SO$_4$, 2 mM MgSO$_4$, 10% (v/v) glycerol, 10 μM ZnCl$_2$, were incubated at 25 °C for a total of 30 min. Pol II was added first and there was a short 5-min incubation between the addition of the template and TFIIB. After addition of 20 μM rCTP, rGTP and rUTP, and 2 μM rATP (Fig. 2c) or 500 μM rCTP, rGTP and rUTP, and 50 μM rATP (Supplementary Fig. 3) supplemented with 5 μCi [α-^{32}P] rATP (Hartmann Analytic) the reactions were incubated at 28 °C for 5 min. The run-off products were separated on a 20% denaturing polyacrylamide gel and visualized by phosphorimaging by a Storm 860 scanner (Molecular Dynamics). Relative RNA levels were quantified with ImageQuant TL (GE Healthcare).

31. Sydow, J. F. *et al.* Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol. Cell* **34,** 710–721 (2009).
32. Berrow, N. S. *et al.* A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res.* **35,** e45 (2007).
33. Sikorski, R. S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae. Genetics* **122,** 19–27 (1989).
34. Seizl, M., Lariviere, L., Pfaffeneder, T., Wenzeck, L. & Cramer, P. Mediator head subcomplex Med11/22 contains a common helix bundle building block with a specific function in transcription initiation complex stabilization. *Nucleic Acids Res.* **39,** 6291–6304 (2011).

LETTER

Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system

Joe Bondy-Denomy[1], April Pawluk[2], Karen L. Maxwell[3] & Alan R. Davidson[1,2]

A widespread system used by bacteria for protection against potentially dangerous foreign DNA molecules consists of the clustered regularly interspaced short palindromic repeats (CRISPR) coupled with *cas* (CRISPR-associated) genes[1]. Similar to RNA interference in eukaryotes[2], these CRISPR/Cas systems use small RNAs for sequence-specific detection and neutralization of invading genomes[3]. Here we describe the first examples of genes that mediate the inhibition of a CRISPR/Cas system. Five distinct 'anti-CRISPR' genes were found in the genomes of bacteriophages infecting *Pseudomonas aeruginosa*. Mutation of the anti-CRISPR gene of a phage rendered it unable to infect bacteria with a functional CRISPR/Cas system, and the addition of the same gene to the genome of a CRISPR/Cas-targeted phage allowed it to evade the CRISPR/Cas system. Phage-encoded anti-CRISPR genes may represent a widespread mechanism for phages to overcome the highly prevalent CRISPR/Cas systems. The existence of anti-CRISPR genes presents new avenues for the elucidation of CRISPR/Cas functional mechanisms and provides new insight into the co-evolution of phages and bacteria.

Predation by phages presents an important challenge to bacterial survival[4], and bacteria have evolved numerous mechanisms to resist phage infection[5]. One such system is the CRISPR/Cas immune system, which is found in 48% of Eubacteria and 95% of Archaea[6]. CRISPR loci contain multiple repeated sequences of approximately 30 base pairs, separated by variable 'spacer' sequences of similar length, which are often identical to segments of phage genomes or other mobile genetic elements[6]. The large single transcript from a CRISPR locus is processed in the repeat regions into small CRISPR RNAs (crRNAs)[7,8] that are complexed with Cas proteins[3,9]. Using the crRNAs as guides, crRNA/Cas complexes cleave foreign DNA molecules at sites bearing complementarity to the crRNAs, thus engendering resistance to phages and other invading DNA molecules[10,11]. Cas proteins also mediate an adaptive function by incorporating short sequences (that is, a new spacer) from newly encountered foreign genomes into CRISPR loci so that these genomes will be destroyed in subsequent encounters[12–14]. The relatively recent elucidation of CRISPR/Cas functions and their obvious similarities to RNA interference (RNAi) systems of eukaryotes have led to vigorous investigation of these systems.

Because phage genes have been discovered that can neutralize most of the prevalent bacterial anti-phage defences[5], the failure to identify genes that counteract the widely occurring CRISPR/Cas systems has been surprising. To search for such 'anti-CRISPR' activity, we investigated the type I–F CRISPR/Cas system[1] of the opportunistic pathogen *Pseudomonas aeruginosa*[14] using a collection of 44 lysogens of *P. aeruginosa* PA14, which each contained a different phage genome (see Methods). In lysogens, phage genomes are integrated into the bacterial genome and are referred to as prophages. Although prophage genes are generally repressed, all prophages have some genes that are actively transcribed. To test whether prophages might express anti-CRISPR activity, we measured the plaquing efficiency of three 'CRISPR-sensitive' phages (JBD18, JBD25 and JBD67) on our collection

of lysogens. The CRISPR-sensitive phages fail to replicate on PA14 owing to the action of the CRISPR/Cas system[14], but are able to replicate on PA14 ΔCR/*cas*, which contains no CRISPR loci or *cas* genes (Fig. 1a and Supplementary Fig. 1a). We identified three lysogenic strains, PA14(JBD24), PA14(MP29) and PA14(JBD30), on which the CRISPR-sensitive phages could form plaques very robustly compared with unlysogenized PA14 (Fig. 1a and Supplementary Fig. 1a). Notably, the plaquing efficiency of the CRISPR-sensitive phages on PA14(JBD30) was equivalent to that on the ΔCR/*cas* strain, indicating that the JBD30 prophage caused complete inactivation of the CRISPR/Cas system. The lower plaquing efficiency of the CRISPR-sensitive phages on the other lysogens relative to their plaquing on the ΔCR/*cas* strain may be due to their production of less potent anti-CRISPR activity. However, these prophages also attenuate plaquing through mechanisms independent of the CRISPR/Cas system, as demonstrated by the partial inhibition of plaquing of the control phage, DMS3, which is not affected by the CRISPR/Cas system (Fig. 1a and Supplementary Fig. 1a).

To assess directly the anti-CRISPR activity of the PA14 lysogens, we used a plasmid-based transformation efficiency assay. The sequences within phages that are targeted by the CRISPR/Cas system are called protospacers. To be targeted, a protospacer sequence must be complementary to a specific spacer sequence in the CRISPR locus and possess a correct protospacer adjacent motif[15,16] (Fig. 1b). Protospacer sequences are named according to the spacer sequence that they match in the PA14 genome (Fig. 1b). We constructed plasmids containing targeted protospacer sequences from phages JBD18 (CRISPR2 locus, spacer 1 or CR2_sp1) and JBD25 (CR1_sp1)[14]. The transformation efficiencies of the plasmids bearing protospacers into unlysogenized PA14 were reduced by at least 90% compared with an empty vector control, whereas no difference in transformation efficiency was seen for the three strains containing prophages expressing anti-CRISPR activity, or for the ΔCR/*cas* strain (Fig. 1c). These data confirm that the prophages isolated in our screen inhibit the PA14 CRISPR/Cas system.

Genome comparisons showed that the phages producing anti-CRISPR activity were closely related to each other, and showed high sequence similarity and synteny with the previously characterized *P. aeruginosa* Mu-like phages D3112, DMS3 and MP22 (Supplementary Table 1). An unusual feature of each of these genomes compared with more distantly related Mu-like phages infecting *P. aeruginosa*[17] and other hosts[18] is the presence of diverse atypical genes at a single position in the region encoding phage head proteins (Fig. 1d and Supplementary Fig. 2). We suspected that some of these genes might encode anti-CRISPR activities; thus, 17 of them were cloned and expressed from a high copy number plasmid in PA14 under the control of an arabinose-inducible promoter. Remarkably, expression of eight of these genes (Supplementary Fig. 3) led to marked increases in the plaquing efficiency of the CRISPR-sensitive phages (Fig. 1e and Supplementary Fig. 1b). Each anti-CRISPR gene allowed these three different phages to evade the CRISPR/Cas system even though they bore distinct protospacer targets (Supplementary Table 1), indicating

[1]Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [2]Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [3]Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada.
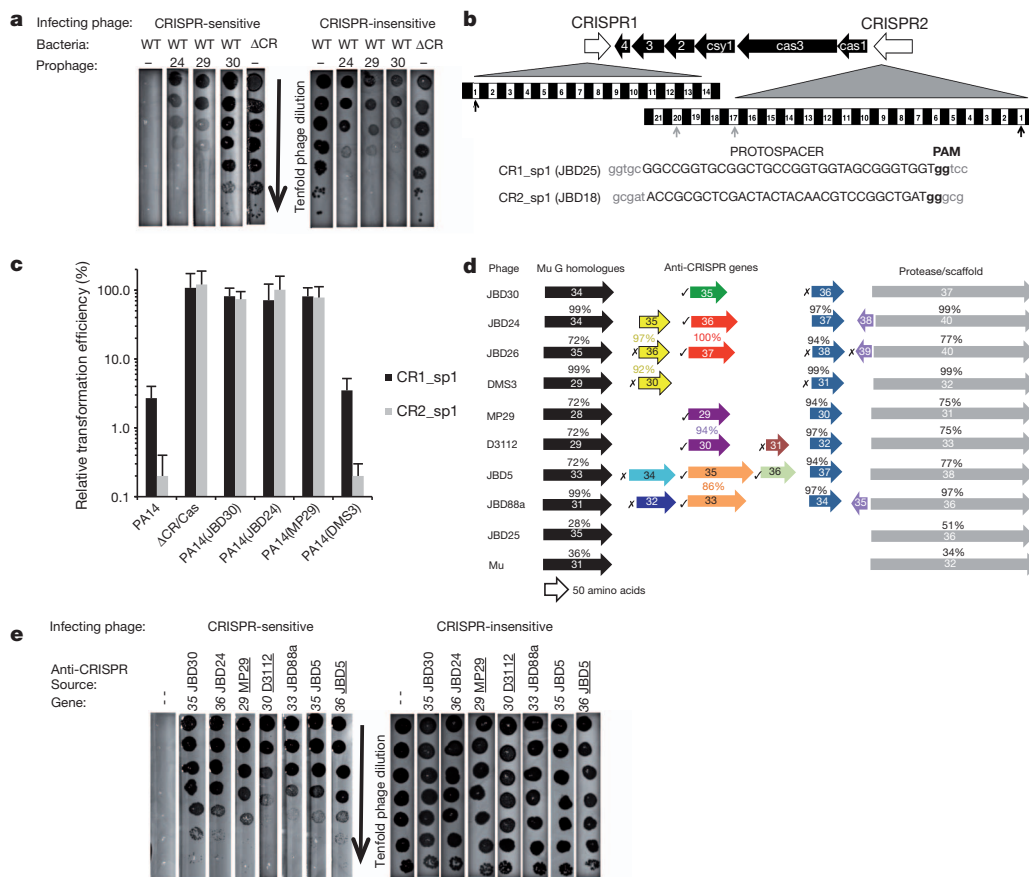
Figure 1 | The CRISPR/Cas system is inhibited by expression of phage genes. a, Tenfold dilutions of lysates of a CRISPR-sensitive phage (JBD18) and a CRISPR-insensitive phage (DMS3) were applied to bacterial lawns of wild-type PA14, PA14 lysogens (JBD24, MP29 or JBD30), and PA14 lacking a CRISPR/Cas system (ΔCR). **b,** Schematic of the PA14 CRISPR loci and *cas* gene region is shown. Expanded versions of each CRISPR locus indicate the number of spacers in each, shown with white boxes, each of which is flanked by repeats denoted by black boxes. Black arrows indicate the CRISPR spacers corresponding to protospacers tested in Fig. 1c and grey arrows indicate the CRISPR spacers corresponding to protospacers tested in Supplementary Fig. 11. The DNA sequences of the protospacers tested in Fig. 1c are shown. **c,** Plasmids containing protospacers shown in Fig. 1b were electroporated into the indicated strains. The relative transformation efficiency was calculated by comparison with the transformation efficiency of the cloning vector containing no protospacer insert. Error bars, s.d. of three biological replicates. **d,** The anti-CRISPR genes of the indicated phages are located in the head gene regions of these genomes between genes homologous to the *G* gene of *E. coli* phage Mu

(black boxes) and genes encoding the protease/scaffold protein of the phage head (grey boxes; see Supplementary Fig. 2). The percentage identities of the proteins encoded by these genes to representatives from JBD30 are shown. The coloured boxes represent putative anti-CRISPR genes. Boxes of the same colour represent closely related genes and the sequence identity of their encoded proteins is indicated. Genes found to mediate anti-CRISPR activity are indicated by tick ('✓') marks and genes tested but showing no anti-CRISPR activity are marked with a cross ('✗') symbol. Unmarked genes were not tested owing to their high similarity to tested genes. The gene box sizes are proportional to the sizes of the proteins in question (scale bar, 50 amino acids), but the spacing of the genes is not to scale. **e,** The same phages from Fig. 1a were tested on PA14 containing empty vector or plasmids expressing the indicated anti-CRISPR genes. Induction of the plasmid promoter with arabinose was required to produce a maximal effect for some of the anti-CRISPR genes as indicated by underlining. The assays shown in Fig. 1a, e are by necessity taken from different plates. The complete plates for these experiments and experiments with the other CRISPR-sensitive phages are shown in Supplementary Fig. 1.

that the anti-CRISPR genes do not function through protection of specific DNA sequences on the phages. We also found that the anti-CRISPR genes can inhibit the type I–F systems found in other *P. aeruginosa* strains, demonstrating that this phenomenon is not particular to strain PA14 (Supplementary Fig. 4). Finally, we found that the anti-CRISPR genes did not inhibit a type I–E CRISPR/Cas system functioning in *Escherichia coli*[19] (Supplementary Fig. 5). Because type I–E is the most closely related CRISPR/Cas system to type I–F[1], we do not expect that these genes would inhibit the function of any of the other more distantly related CRISPR/Cas systems.

For the type I–F CRISPR/Cas system to function, transcription of pre-crRNA from the CRISPR locus must occur, followed by processing into small crRNAs and incorporation of these RNAs into a complex with Cas proteins[3,7]. The stable maintenance of processed crRNA within the cell requires Cas proteins. Thus, the lack of any *P. aeruginosa* Cas protein except Cas1, which is involved in spacer acquisition, causes a marked reduction in crRNA amounts[20]. As can be seen in

Supplementary Fig. 6a, expression of five different anti-CRISPR genes in PA14 caused no change in the amount of processed crRNA molecules as detected by northern blotting. The normal amount of crRNA observed implies that expression of anti-CRISPR genes does not cause a reduction in the expression amounts of the CRISPR loci or *cas* genes. Supporting this finding, we also showed that transcription levels of the *cas* genes *cas3* and *csy3* were unaffected by the anti-CRISPR genes as assessed by reverse transcriptase quantitative PCR (RT–qPCR) (Supplementary Fig. 6b, c). Furthermore, β-galactosidase activity produced from a chromosomally located *csy3::lacZ* fusion gene was not perturbed by expression of any of the anti-CRISPR genes (Supplementary Fig. 6d). We conclude from these experiments that the anti-CRISPR genes exert their effects at a step occurring after formation of the crRNA-Cas complex, and that there is no effect on biogenesis of either the crRNA or Cas proteins.

Despite the common genomic positions of the anti-CRISPR genes in very similar *P. aeruginosa* phages, these eight genes are predicted to

encode five different proteins with completely distinct sequences (Supplementary Fig. 3). Sequence similarity searches with each of these proteins yielded fewer than 15 different significant hits in total, of which all but four were proteins encoded in genomes of closely related phages or prophages (Supplementary Table 2). One of the non-phage-associated anti-CRISPR protein homologues, which is 43% identical to the product of gene *33* of phage JBD88a (JBD88a gp33), is encoded in an active pathogenicity island of a highly virulent *P. aeruginosa* clinical isolate that is probably transferred by conjugation between *P. aeruginosa* strains[21]. This island contains four protospacers with correct protospacer adjacent motifs and 100% identity to CRISPR spacers in various *P. aeruginosa* strains[22]. The three other non-phage-associated anti-CRISPR homologues are also found in regions of *Pseudomonas* genomes that may be mobile elements as indicated by presence of genes in these regions encoding homologues of proteins involved in DNA transfer and/or type IV secretion (Supplementary Fig. 7). Thus, these putative bacterial anti-CRISPR genes may increase the fitness for inter-strain transfer of these mobile elements by inactivating the CRISPR/Cas system of a recipient strain.

Because the crRNA/Cas complex is guided by RNA, anti-CRISPR activity might be mediated by a non-coding RNA molecule or a protein encoded by an anti-CRISPR gene. We addressed this issue by performing experiments on JBD30 gene *35*. A nonsense mutation at the third codon and two different frameshift mutations were introduced to the plasmid encoding gene *35*. Each of these mutations abrogated anti-CRISPR activity (Supplementary Fig. 8), implying that translation of this region was required for function. Because these experiments did not rule out a combined role for anti-CRISPR non-coding RNA and protein, two variant genes were synthesized that encoded the same amino-acid sequence as gene *35*, yet had DNA sequences that differed by approximately 35% through variation of codon wobble positions (Supplementary Fig. 9). As shown in Supplementary Fig. 8, each of these synthetic versions of gene *35* imparted full anti-CRISPR activity. These data demonstrate that anti-CRISPR protein is required for anti-CRISPR activity and that a direct mechanistic role for a gene *35*-encoded RNA is unlikely.

The genomes of six of the seven 'anti-CRISPR phages' (that is, those phages bearing active anti-CRISPR genes, Fig. 1d) contain at least one functional protospacer (Supplementary Table 1); thus, their replication should be inhibited by the PA14 CRISPR/Cas system. However, each was able to form plaques on PA14 with 100% efficiency compared with the ΔCR/*cas* strain (Supplementary Fig. 10). Using the transformation efficiency assay, we confirmed that the two protospacers found most commonly in the anti-CRISPR phages were indeed targeted by the PA14 CRISPR/Cas system (Supplementary Fig. 11). These results implied that the anti-CRISPR phages were able to replicate on PA14 because they possess anti-CRISPR genes. To address this hypothesis, a frameshift mutation was introduced into the phage JBD30 anti-CRISPR gene (gene *35*). This mutant phage was unable to replicate on wild-type PA14 but still replicated robustly on the ΔCR/*cas* strain, demonstrating the requirement of the anti-CRISPR gene for replication in cells bearing an intact CRISPR/Cas system (Fig. 2a). To determine whether the introduction of an anti-CRISPR gene into a CRISPR-sensitive phage would allow that phage to evade CRISPR/Cas immunity, we used a CRISPR-sensitive mutant of phage DMS3, called DMS3*m*[14]. This phage possesses a functional protospacer and is very similar in sequence to the anti-CRISPR phages, yet it contains no functional anti-CRISPR gene (Fig. 1c). Taking advantage of the high DNA sequence identity between phages DMS3*m* and JBD30, *in vivo* homologous recombination was used to create a version of DMS3*m* bearing JBD30 gene *35* (Fig. 2b, Supplementary Fig. 12 and Methods). As shown in Fig. 2c, the introduction of this gene into DMS3*m* resulted in a 10^6-fold increase in plaquing efficiency on PA14, clearly demonstrating that an anti-CRISPR gene present on an infecting phage allows that phage to overcome the CRISPR/Cas system. By testing lysogens of the DMS3*m* and JBD30 mutant prophages we found that
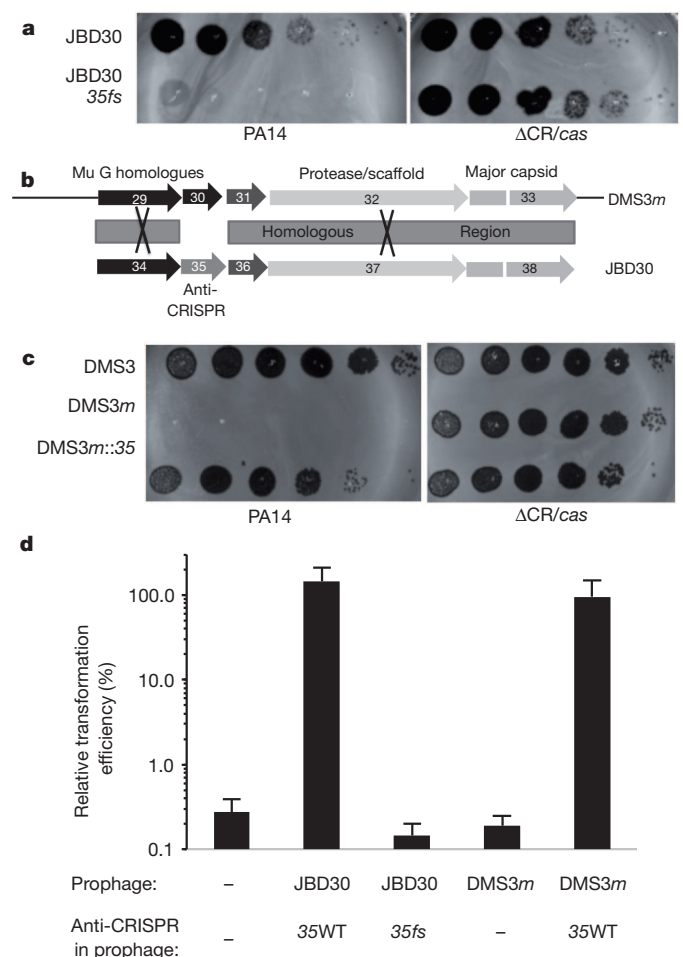

Figure 2 | An anti-CRISPR gene protects phages from the CRISPR/Cas system during infection. a, Tenfold dilutions of lysates of anti-CRISPR phage JBD30, and the same phage with a frameshift mutation introduced into the anti-CRISPR gene *35* (gene *35fs*), were applied to lawns of PA14 or PA14 ΔCR/*cas*. This experiment was done in a manner similar to those shown in Fig. 1a. **b**, Schematic representation of the *in vivo* homologous recombination between phage DMS3*m* and the anti-CRISPR region from phage JBD30. The 'X' approximately marks the mapped region of recombination, up- and downstream of the anti-CRISPR gene *35* from JBD30 with details shown in Supplementary Fig. 12 and Methods. **c**, Tenfold dilutions of lysates of a CRISPR-insensitive phage (DMS3), a CRISPR-sensitive phage (DMS3*m*) and DMS3*m* with anti-CRISPR gene *35* from JBD30 inserted (DMS3*m*::gene *35*) were applied to lawns of PA14 or PA14 ΔCR/*cas*. **d**, A plasmid containing a protospacer matching CR1_sp1 (shown in Fig. 1b) was electroporated into the indicated lysogens or parent strain. As indicated, the prophages within these lysogens contain either a wild-type version of anti-CRISPR gene *35*, a frameshift mutant of this gene or no anti-CRISPR gene.

a protospacer-bearing plasmid transformed cells efficiently only those lysogens in which an intact anti-CRISPR gene was present (Fig. 2d). These assays demonstrate the necessity and sufficiency of the anti-CRISPR gene for inhibition of the CRISPR/Cas system.

The adaptive nature of the CRISPR/Cas system and the widespread occurrence of CRISPR regions in bacterial genomes suggest that this system could be the most powerful weapon possessed by bacteria to resist invasion by foreign DNA. Before our work, the only known mechanism for phages to evade CRISPR/Cas systems was by mutation[12,23], which is a low-frequency event. Here we have provided the first demonstration that the *in vivo* activity of a CRISPR/Cas system is profoundly inhibited by any one of five different 'anti-CRISPR' genes. The existence of anti-CRISPR genes is one possible explanation of how phages have continued to proliferate despite the ubiquity and potency of CRISPR/Cas systems. The possibility that anti-CRISPR genes are

diverse and widespread among phages and other mobile genetic elements may account for the large diversity of CRISPR/Cas systems and the existence of multiple CRISPR/Cas system types within single bacterial strains. This proliferation of CRISPR/Cas systems may be driven by the concomitant proliferation and diversification of anti-CRISPR genes. This newly discovered arms race may have a profound effect on the evolution of both phage and bacterial genomes, and knowledge of anti-CRISPR genes will be crucial for understanding this process. The failure to detect anti-CRISPR genes until now may be due only to a lack of systematic searches using naturally functioning *in vivo* systems. Future studies to discover more anti-CRISPR genes and explain the mechanisms of their inhibition of CRISPR/Cas systems will provide new inroads for illumination of CRISPR/Cas function.

METHODS SUMMARY

Phage plaque formation assays. Overnight cultures of *P. aeruginosa* were inoculated into molten soft agar and poured onto lysogeny broth agar plates. Once solidified, 3 μl aliquots of phage lysates at tenfold dilutions were pipetted onto the agar. Plates were incubated overnight at 30 °C or 37 °C.

Lysogen construction. Phages were isolated from clinical and environmental isolates through induction with mitomycin C and subsequent plaque purification on a susceptible host. PA14 lysogens were isolated by pipetting phage lysates onto a PA14 lawn, created as described above, followed by isolation of individual colonies from the zone of clearing. Colonies were assayed for the production of the correct phage to confirm lysogeny.

Transformation efficiency assay. To construct plasmids containing protospacer sequences that are susceptible to the PA14 CRISPR/Cas system, two 42-base complementary synthetic oligonucleotides containing the protospacer were ligated to a plasmid vector, pHERD30T.

Full Methods and any associated references are available in the online version of the paper.

1. Makarova, K. S. *et al.* Evolution and classification of the CRISPR-Cas systems. *Nature Rev. Microbiol.* **9,** 467–477 (2011).
2. Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* **1,** 7 (2006).
3. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321,** 960–964 (2008).
4. Rodriguez-Valera, F. *et al.* Explaining microbial population genomics through phage predation. *Nature Rev. Microbiol.* **7,** 828–836 (2009).
5. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nature Rev. Microbiol.* **8,** 317–327 (2010).
6. Jore, M. M., Brouns, S. J. J. & van der Oost, J. RNA in defense: CRISPRs protect prokaryotes against mobile genetic elements. *Cold Spring Harb. Perspect. Biol.* **4,** http://dx.doi.org/10.1101/cshperspect.a003657 (2012).
7. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J. A. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329,** 1355–1358 (2010).
8. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471,** 602–607 (2011).
9. Wiedenheft, B. *et al.* Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* **477,** 486–489 (2011).
10. Westra, E. R. *et al.* CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol. Cell* **46,** 595–605 (2012).
11. Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468,** 67–71 (2010).
12. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315,** 1709–1712 (2007).
13. Yosef, I., Goren, M. G. & Qimron, U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli. Nucleic Acids Res.* **40,** 5569–5576 (2012).
14. Cady, K. C., Bondy-Denomy, J., Heussler, G. E., Davidson, A. R. & O'Toole, G. A. The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J. Bacteriol.* **194,** 5728–5738 (2012).
15. Marraffini, L. A. & Sontheimer, E. J. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463,** 568–571 (2010).
16. Mojica, F. J. M., Diez-Villasenor, C., Garcia-Martinez, J. & Cano, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155,** 733–740 (2009).
17. Braid, M. D., Silhavy, J. L., Kitts, C. L., Cano, R. J. & Howe, M. M. Complete genomic sequence of bacteriophage B3, a Mu-like phage of *Pseudomonas aeruginosa. J. Bacteriol.* **186,** 6560–6574 (2004).
18. Morgan, G. J., Hatfull, G. F., Casjens, S. & Hendrix, R. W. Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in *Haemophilus, Neisseria* and *Deinococcus. J. Mol. Biol.* **317,** 337–359 (2002).
19. Datsenko, K. A. *et al.* Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nature Commun.* **3,** 945 (2012).
20. Cady, K. C. & O'Toole, G. A. Non-identity-mediated CRISPR-bacteriophage interaction mediated via the Csy and Cas3 proteins. *J. Bacteriol.* **193,** 3433–3445 (2011).
21. Battle, S. E., Meyer, F., Rello, J., Kung, V. L. & Hauser, A. R. Hybrid pathogenicity island PAGI-5 contributes to the highly virulent phenotype of a *Pseudomonas aeruginosa* isolate in mammals. *J. Bacteriol.* **190,** 7130–7140 (2008).
22. Cady, K. C. *et al.* Prevalence, conservation and functional analysis of *Yersinia* and *Escherichia* CRISPR regions in clinical *Pseudomonas aeruginosa* isolates. *Microbiology* **157,** 430–437 (2011).
23. Semenova, E. *et al.* Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl Acad. Sci. USA* **108,** 10098–10103 (2011).

Author Contributions J.B.-D. designed experiments, performed experiments and wrote the manuscript, A.P. performed experiments, K.L.M. supervised experiments, and A.R.D. designed experiments and wrote the manuscript.

Author Information Phage genomes have been deposited in the National Center for Biotechnology Information under accession numbers JX434030 (JBD5), JX434031 (JBD24), JX434032 (JBD30) and JX434033 (JBD88a). Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.R.D. (alan.davidson@utoronto.ca).

METHODS

Strains and growth conditions. All bacterial strains, phages, plasmids and primers used in this study are listed in Supplementary Table 3 (File 1). *Pseudomonas aeruginosa* UCBPP-PA14 (PA14), other *P. aeruginosa* isolates and *E. coli* DH5α were grown on lysogeny broth agar or liquid medium at 37 °C. Lysogeny broth was supplemented with gentamicin (50 µg ml^{-1} for *P. aeruginosa* and 15 µg ml^{-1} for *E. coli*) to maintain the pHERD30T plasmid or ampicillin (100 µg ml^{-1}) for *E. coli* and carbenicillin (300 µg ml^{-1}) for *P. aeruginosa* with the pHERD20T plasmid.

Phage induction and isolation. Phages were isolated from a diverse panel of 88 clinical and environmental isolates of *P. aeruginosa* by inducing the resident prophages in these strains. The strains were grown at 37 °C to early log phase ($D_{600\,nm} = 0.5$) and the DNA-damaging agent mitomycin C (3 µg ml^{-1}) was added to the culture to induce prophages. Treatment was allowed to continue until lysis was visible (approximately 4–5 h) and chloroform was added for 15 min. The lysate was centrifuged at 10,000*g* for 10 min and the supernatant was kept and stored over 100 µl of chloroform at 4 °C. These lysates were subjected to plaque assays on bacterial lawns of 20 indicator strains from the initial collection. Visible plaques indicated the presence of phage originating from a given lysate. All phages were subjected to three rounds of plaque purification before further characterization.

Plaque assays. Plaque assays were conducted at 30 °C or 37 °C on lysogeny broth agar (1.5%) plates with lysogeny broth top agar (0.7%), both supplemented with MgSO$_4$ (10 mM). One hundred and fifty microlitres of an overnight culture were mixed with molten soft agar and poured onto lysogeny broth agar plates, onto which 3 µl aliquots of phage suspensions were spotted. The observed circular zones of clearing (plaques) were due to phage replication and lysis of the bacterial host. Alternatively, full plate assays were conducted where phage and bacteria were mixed, pre-adsorbed for 15 min at 37 °C before adding to top agar and pouring. In the images shown in this paper, 3 µl of phage were spotted in each example and the most concentrated spot contained approximately 5×10^6 to 1×10^7 plaque-forming units. For the preparation of highly pure phage stocks, phage were precipitated with 10% (w/v) polyethylene glycol (PEG) 8000 and subjected to two caesium chloride equilibrium centrifugation gradients. These preparations were used for DNA extraction and subsequent sequencing.

Lysogen construction. Lysogens were constructed by spotting serial dilutions of phage lysate on strain PA14 and streaking out bacteria (that is, putative lysogens) from the inside of a clearing resulting from a cluster of plaques. Colonies were then screened to confirm resistance to the phage used to lysogenize the strain. The putative lysogens were grown in liquid culture, and the presence of spontaneously produced phage in the supernatant that could plaque on the original wild-type strain confirmed lysogeny. All PA14 lysogens were constructed with phages isolated from this work (i.e. JBD phages), with the exception of phages DMS3 (ref. 24), MP22 (ref. 25), MP29 (ref. 26) and D3112 (ref. 27).

DNA extraction/phage genome sequencing. Phage genomic DNA was extracted for sequencing from 500 µl of a caesium-chloride-purified phage preparation with a titre between 10^{11} and 10^{12} plaque-forming units per millilitre. Phages were disrupted and proteolysed by treatment with EDTA (20 mM), proteinase K (50 µg ml^{-1}) and SDS (0.5% w/v) at 56 °C for 1 h, then the DNA was extracted with phenol/chloroform and precipitated with ethanol. Phage genomes were sequenced with either Illumina Solexa or 454 high-throughput sequencing. Accession numbers for phage genomes are provided in Supplementary Table 3.

Phage genome analysis. New phage genome sequences were first analysed using BLASTn to assess general similarity to previously sequenced phages. To predict open reading frames and align multiple phages for comparison, the RAST program[28] was used. Comparison and analysis of specific phage proteins was done using RAST, CLUSTAL[29] or BLASTp/PSI-BLAST[30].

Plasmid construction. A shuttle vector that replicates in *E. coli* and *P. aeruginosa*, pHERD30T[31], was used for cloning and expression of genes in *P. aeruginosa*. This vector has an arabinose-inducible promoter and a selectable gentamicin marker. A related plasmid, pHERD20T, was used for the anti-CRISPR knockout and knock-in experiments, which was ampicillin/carbenicillin selectable. Inserts were amplified by PCR. Vector and insert were digested with the indicated restriction enzymes (Supplementary Table 3), ligated, and the ligation mix was used to transform *E. coli* DH5α. All plasmid constructs were verified by sequencing using primers that annealed to sites outside the multiple cloning site. Other mutations to JBD30 gene *35* were introduced using primers shown in Supplementary Table 3.

To produce versions of JBD30 gene *35* with divergent DNA sequences (JBD30–35 varA and varB), sequences containing maximal numbers of silent mutations were designed manually and synthesized by GenScript USA. The synthesized genes were subcloned into pHERD30T by digestion with NcoI and HindIII.

Transformation efficiency assays. Oligonucleotides were synthesized to match a desired 32-base protospacer (Supplementary Table 3) with the upstream and downstream five bases. Extra bases were added on the 5′ and 3′ ends of each oligonucleotide to produce NcoI and HindIII sticky ends after annealing. The

oligonucleotides were annealed and ligated into the pHERD30T shuttle vector. For transformation assays into *P. aeruginosa*, standard electroporation protocols were used. One millilitre of an overnight culture was washed twice in 300 mM sucrose and concentrated tenfold. Subsequently, 300–500 ng of plasmid were electroporated, then the cells were incubated in antibiotic-free lysogeny broth medium for 1 h at 37 °C. Cells were plated on gentamicin (50 µg ml^{-1}) selective media and colonies were counted after overnight growth at 37 °C. The number of colonies was normalized to the dilution factor and the mass of plasmid transformed to yield transformants per microgram of DNA. The relative transformation efficiency for each protospacer was calculated as a percentage of the transformation efficiency obtained for the empty vector control.

In assays shown in Fig. 2d, the parental strain in all cases was PA14 lacking the CRISPR2 locus owing to CRISPR2-mediated inhibition, which would prevent lysogeny of phages JBD30 and DMS3*m* lacking a functional anti-CRISPR. This was necessary so that phage JBD30 lacking its anti-CRISPR gene and phage DMS3*m* would be able to form lysogens. As shown in the left-most column, the transformation efficiency of the construct bearing CR1_sp1 with PA14 ΔCRISPR2 was approximately tenfold less than the same construct in transforming wild-type PA14 (Fig. 1c). This was a consistent result, presumably because of the increased expression of crRNA from the CRISPR1 locus in the absence of CRISPR2, although this has not been explored.

DMS3 recombination. Wild-type phage DMS3 contains a protospacer region in gene *42* with five mismatches to the 32 nt CR2_sp1 crRNA produced by PA14 (ref. 20). A constructed mutant DMS3 phage (DMS3*m*) described previously[14] as DMS3$_{100\%}$, contains five point mutations in gene *42* creating a 100% match to this crRNA. Owing to targeting by the CRISPR/Cas system, this mutant phage is unable to form plaques on wild-type PA14, with the exception of rare (fewer than 10^{-6}) escaper mutants. Despite overall genomic similarity between phages DMS3*m* and JBD30, phage DMS3*m* does not contain an anti-CRISPR gene of its own. Therefore, DMS3*m* was used to generate a recombinant phage containing the anti-CRISPR gene from JBD30 (gene *35*). Cells containing a pHERD20T construct containing JBD30 genes *34–38* (that is, the anti-CRISPR gene from JBD30 with large flanking regions) were infected with DMS3*m* and recombinant phages were selected by plating on wild-type PA14. Individual plaques were picked and purified three times by re-plating on PA14, which ultimately led to the isolation of three independent recombinants. Preliminary screening involved PCR reactions with primers to the anti-CRISPR gene as well as primers to the protospacer region (gene *42*) which were specific for a protospacer with 100% identity to the CR2_sp1 crRNA. Putative recombinant phages were then subject to three PCR reactions and subsequent sequencing. The protospacer containing region (gene *42*) was sequenced to confirm that the recombinants had maintained 100% complementarity with the CR2_sp1 crRNA. The expected site of anti-CRISPR gene recombination (between DMS3 genes *29* and *31*) was amplified using primers matching phage DMS3, outside the region that was on the plasmid used for recombination, and this was analysed to identify the site of recombination (shown in Supplementary Fig. 12). The anti-CRISPR gene was also sequenced in recombinant phages to confirm that it was present and in-frame in all three recombinants (not shown). These reactions confirmed that the anti-CRISPR gene had been acquired and these phages were used in the experiments described here (Fig. 2c, d and Supplementary Fig. 12). Escapers were also identified with mutations in the protospacer and no recombination identified.

Northern blot. Northern blots were conducted as described previously[20], with exceptions described below. Wild-type PA14, PA14Δ*csy4*, PA14Δ*cr1/cr2* (both CRISPR loci deleted, *cas* genes intact) and PA14 lysogens were grown in lysogeny broth. Total RNA was extracted from log-phase cultures ($D_{600nm} = 0.8$) using the *mir*Vana microRNA isolation kit (Ambion) and 5 µg total RNA were run on a 10% TBE-urea polyacrylamide gel and stained with SYBR Gold Nucleic Acid Gel Stain (Invitrogen) before transferring RNA to a nylon membrane at 200 mA for 1 h. A radiolabelled probe corresponding to the last four spacers and three repeats of the PA14 CRISPR locus was used[20]. Prehybridization (blocking) was conducted using 50% formamide, 5× Denhardts, 0.5% SDS, 6× SSC and 100 µg ml^{-1} single-stranded DNA at 42 °C for 2 h. Probing was conducted at 42 °C for 16 h using fresh prehybridization buffer, but with the single-stranded DNA omitted and radiolabelled probe added. Wash solution 1 consisted of 2× SSC and 1% SDS, whereas wash solution 2 consisted of 0.2× SSC and 0.1% SDS. Wash solution 1 was used for two 10 min washes at 25 °C, two 30 min washes at 65 °C, and wash solution 2 for one 10 min wash at 25 °C. Blots were developed using a phosphor screen and imager. A low-range single-stranded RNA ladder (NEB) was also used to confirm the location of 5S RNA and crRNA.

RT–qPCR. Total RNA extracts were treated with DNase (Ambion) to remove DNA and 1 ng of total RNA was used in a series of RT–qPCR reactions. Reactions were conducted in an Eppendorf qPCR cycler, using VWR white plates with the SensiFAST No-ROX One-step Kit (Bioline). For absolute quantification, a PCR

reaction was conducted, amplifying genomic DNA with 'external' primers. This product was gel extracted, quantified and diluted to generate a standard curve. All RT–qPCR reactions were done with 'internal' primers that were designed to anneal inside the external primers for the purified PCR product. A housekeeping gene, rpsL, was used for relative quantification. For RT–qPCR reactions, 1 ng of total RNA was used in each reaction, performed in duplicate. Reverse transcription was conducted using a gene-specific primer to generate complementary DNA in a one-step reaction. The lack of contaminating DNA was confirmed by inclusion of controls for each sample without reverse transcriptase added.

β-Galactosidase assays. Overnight cultures were subcultured 1:100 into lysogeny broth containing 0.1% arabinose to induce anti-CRISPR gene expression from the pHERD30T plasmid, then grown at 37 °C to an attenuance ($D_{600\,nm}$) of 0.3–0.6. Cultures were diluted 1:1 in complete Z buffer, in triplicate. Two drops each of 0.1% SDS and chloroform were added to each sample and, after vortexing, 200 µl ONPG was added. Samples were vortexed to begin the reaction, then incubated at 30 °C without shaking for 20–30 min. Absorbance measurements were taken at 420 and 550 nm, and β-galactosidase activity was calculated for each technical replicate using the Miller equation. Data are expressed relative to cells containing the pHERD30T empty vector for at least five biological replicates.

Anti-CRISPR knockout. Three independent PCR products were generated which contained overhangs, which facilitated the generation of a pHERD20T (Amp^R, 20T) plasmid with the region from phage JBD30 that lies up and downstream of the anti-CRISPR (gene 35) with an inserted gentamicin (Gm) resistance cassette and the first 60 base pairs of gene 35 absent. Primers used are outlined in Supplementary Table 3. Product 1: gene 34 with 5′ overhangs to 20T, 3′ overhangs to Gm. Product 2: Gm cassette with 5′ overhangs to Gene 34, 3′ overhangs to Gene 35 (lacking first 60 base pairs). Product 3: gene 35 (lacking first 60 base pairs) with 5′ overhangs to Gm, 3′ overhangs to 20T.

All three PCR products were gel purified and amplified together in a reaction with Vent DNA Polymerase using primers pHERD-34-F and pHERD-36-R. This reaction yielded a DNA product that was gel extracted and re-amplified with Taq DNA polymerase. The pHERD20T vector was digested with EcoRI and HindIII and then infused using the IN-Fusion EcoDry cloning kit (Clontech) following the manufacturer's instructions. The mixture was used to transform E. coli DH5α using standard protocols with recovery in SOC followed by plating and selection with ampicillin. Plasmids were isolated from colonies that were both gentamicin- and ampicillin-resistant and confirmed by sequencing and diagnostic restriction enzyme digestion. The plasmid containing a gentamicin resistance cassette, flanked by regions of homology to the 300 base pairs upstream and downstream of the coding region of JBD30 gene 35 was electroporated into a PA14ΔCR/cas(JBD30) lysogen to generate recombinant phages. Isolated gentamicin resistant colonies were grown in liquid culture to allow phage production, and a cell-free phage preparation from the supernatant was used to infect ΔCR/cas cells overnight in liquid culture. Surviving cells were plated on gentamicin to select for newly formed ΔCR(JBD30) lysogens with a prophage containing the gentamicin insert in gene 35. Resulting colonies were confirmed to be gentamicin resistant and carbenicillin sensitive (confirming no plasmid transfer had occurred). The correct insertion of the gentamicin cassette in the phage genome was confirmed by PCR and DNA sequencing. The presence of the gentamicin cassette in the phage created plaques that were smaller, more turbid and required prolonged incubation. To remove the cassette, pHERD20T containing JBD30 genes 34–38 with

either wild-type sequence or a frameshift introduced in gene 35 was electroporated into the knockout strain under carbenicillin selection. Single colonies were grown in liquid and plaque assays conducted with the supernatant to identify recombinants. Plaques with wild-type morphology and growth kinetics were picked, purified and confirmed by PCR and sequencing to contain either the wild-type sequence or a frameshift mutation in gene 35 corresponding to the input vector.

The frameshift mutation in JBD30 gene 35 was introduced into pHERD20T containing JBD30 genes 34–38 with a PCR reaction with primers 30–35fsF/R. Eighteen cycles were conducted with 9 min extension time using Pfu DNA polymerase and subsequent DpnI digest to remove parental plasmid.

Type I–E CRISPR/Cas system assays. One and a half millilitres of an overnight culture was washed twice with 1 ml ice-cold sterile water, washed once with 1 ml ice-cold 10% glycerol, and re-suspended in 100 µl ice-cold 10% glycerol. Approximately 300 ng of plasmid was electroporated and the cells recovered. Cells were then centrifuged, re-suspended in 100 µl lysogeny broth medium, and plated on lysogeny broth agar plates containing 15 µg ml^{-1} gentamicin. Plates were grown overnight at 37 °C. Overnight cultures of E. coli BW40114 and BW40119 (ref. 19) containing the empty vector or anti-CRISPR genes were diluted 100-fold into lysogeny broth medium containing 15 µg ml^{-1} gentamicin and grown at 37 °C for 4 h. Isopropyl-β-D-thiogalactoside (IPTG) and arabinose were then added to a final concentration of 1 mM each. The cultures were grown under induction for 2 h, then 1.5 ml of each culture was concentrated into 100 µl and mixed with 3 ml molten lysogeny broth soft agar. The mixture was poured onto thick lysogeny broth agar plates containing 15 µg ml^{-1} gentamicin, 1 mM IPTG and 1 mM arabinose. Tenfold serial dilutions of M13 phage lysate were spotted onto the plates and incubated overnight at 30 °C. Lysates which were produced from targeting cells (BW40119) containing anti-CRISPR constructs or empty vector under conditions in which the plasmid promoter was repressed (−, 0.2% glucose) or induced (+, 1 mM arabinose) and analysed by SDS–PAGE to assess the expression of anti-CRISPR proteins.

24. Zegans, M. E. et al. Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of Pseudomonas aeruginosa. J. Bacteriol. **191,** 210–219 (2009).
25. Heo, Y.-J., Chung, I.-Y., Choi, K. B., Lau, G. W. & Cho, Y.-H. Genome sequence comparison and superinfection between two related Pseudomonas aeruginosa phages, D3112 and MP22. Microbiology **153,** 2885–2895 (2007).
26. Chung, I.-Y. & Cho, Y.-H. Complete genome sequences of two Pseudomonas aeruginosa temperate phages, MP29 and MP42, which lack the phage-host CRISPR interaction. J. Virol. **86,** 8336 (2012).
27. Wang, P. W., Chu, L. & Guttman, D. S. Complete sequence and evolutionary genomic analysis of the Pseudomonas aeruginosa transposable bacteriophage D3112. J. Bacteriol. **186,** 400–410 (2004).
28. Aziz, R. K. et al. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9,** 75 (2008).
29. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. Bioinformatics **23,** 2947–2948 (2007).
30. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. J. Mol. Biol. **215,** 403–410 (1990).
31. Qiu, D., Damron, F. H., Mima, T., Schweizer, H. P. & Yu, H. D. PBAD-based shuttle vectors for functional analysis of toxic and highly regulated genes in Pseudomonas and Burkholderia spp. and other bacteria. Appl. Environ. Microbiol. **74,** 7422–7426 (2008).

LETTER

Mosaic *PPM1D* mutations are associated with predisposition to breast and ovarian cancer

Elise Ruark[1]*, Katie Snape[1]*, Peter Humburg[2]*, Chey Loveday[1], Ilirjana Bajrami[3], Rachel Brough[3,4], Daniel Nava Rodrigues[3], Anthony Renwick[1], Sheila Seal[1], Emma Ramsay[1], Silvana Del Vecchio Duarte[1], Manuel A. Rivas[2,5], Margaret Warren-Perry[1], Anna Zachariou[1], Adriana Campion-Flora[3], Sandra Hanks[1], Anne Murray[1], Naser Ansari Pour[1], Jenny Douglas[1], Lorna Gregory[2], Andrew Rimmer[2], Neil M. Walker[6], Tsun-Po Yang[7], Julian W. Adlard[8], Julian Barwell[9], Jonathan Berg[10], Angela F. Brady[11], Carole Brewer[12], Glen Brice[13], Cyril Chapman[14], Jackie Cook[15], Rosemarie Davidson[16], Alan Donaldson[17], Fiona Douglas[18], Diana Eccles[19], D. Gareth Evans[20], Lynn Greenhalgh[21], Alex Henderson[18], Louise Izatt[22], Ajith Kumar[23], Fiona Lalloo[24], Zosia Miedzybrodzka[25], Patrick J. Morrison[26], Joan Paterson[27], Mary Porteous[28], Mark T. Rogers[29], Susan Shanley[30], Lisa Walker[31], Martin Gore[32], Richard Houlston[1], Matthew A. Brown[33], Mark J. Caulfield[34], Panagiotis Deloukas[7], Mark I. McCarthy[2,35,36], John A. Todd[6], The Breast and Ovarian Cancer Susceptibility Collaboration†, Wellcome Trust Case Control Consortium†, Clare Turnbull[1,30], Jorge S. Reis-Filho[3], Alan Ashworth[3], Antonis C. Antoniou[37], Christopher J. Lord[3], Peter Donnelly[2,38] & Nazneen Rahman[1,30]

Improved sequencing technologies offer unprecedented opportunities for investigating the role of rare genetic variation in common disease. However, there are considerable challenges with respect to study design, data analysis and replication[1]. Using pooled next-generation sequencing of 507 genes implicated in the repair of DNA in 1,150 samples, an analytical strategy focused on protein-truncating variants (PTVs) and a large-scale sequencing case–control replication experiment in 13,642 individuals, here we show that rare PTVs in the p53-inducible protein phosphatase *PPM1D* are associated with predisposition to breast cancer and ovarian cancer. *PPM1D* PTV mutations were present in 25 out of 7,781 cases versus 1 out of 5,861 controls ($P = 1.12 \times 10^{-5}$), including 18 mutations in 6,912 individuals with breast cancer ($P = 2.42 \times 10^{-4}$) and 12 mutations in 1,121 individuals with ovarian cancer ($P = 3.10 \times 10^{-9}$). Notably, all of the identified *PPM1D* PTVs were mosaic in lymphocyte DNA and clustered within a 370-base-pair region in the final exon of the gene, carboxy-terminal to the phosphatase catalytic domain. Functional studies demonstrate that the mutations result in enhanced suppression of p53 in response to ionizing radiation exposure, suggesting that the mutant alleles encode hyperactive PPM1D isoforms. Thus, although the mutations cause premature protein truncation, they do not result in the simple loss-of-function effect typically associated with this class of variant, but instead probably have a gain-of-function effect. Our results have implications for the detection and management of breast and ovarian cancer risk. More generally, these data provide new insights into the role of rare and of mosaic genetic variants in common conditions, and the use of sequencing in their identification.

There is strong evidence that rare genetic variation is important in breast and ovarian cancer predisposition[2,3]. In the 1990s, genome-wide linkage analysis and positional cloning led to the identification of the DNA repair genes *BRCA1* and *BRCA2*, rare mutations of which confer substantial risks of both diseases[2,3]. More recently, through case-control resequencing studies of candidate genes we, and others, have discovered rare variants that confer moderate risks of breast and/or ovarian cancer[4–10]. These cancers are therefore exemplars of the rare variant, common disease hypothesis.

The successful studies so far have focused on genes encoding proteins involved in DNA repair such as *PALB2*, *ATM*, *CHEK2*, *BRIP1*, *RAD51C* and *RAD51D* (refs 4–10). These genes are characterized by several, very rare, loss-of-function mutations, usually PTVs, which predispose carriers to breast and/or ovarian cancer[4–10]. To investigate the role of DNA repair genes in cancer susceptibility further, we sequenced 507 genes (the 'DNA repair panel') in 1,150 individuals with breast cancer from the United Kingdom, 69 of whom also had ovarian cancer (Supplementary Tables 1 and 2 and Supplementary

[1]Division of Genetics & Epidemiology, The Institute of Cancer Research, Sutton SM2 5NG, UK. [2]The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK. [3]The Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London SW3 6JB, UK. [4]Cancer Research UK Gene Function Laboratory, The Institute of Cancer Research, London SW3 6JB, UK. [5]Nuffield Department of Clinical Medicine, University of Oxford, Oxford OX3 7LD, UK. [6]Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research, University of Cambridge, Addenbrooke's Hospital, Cambridge CB2 0XY, UK. [7]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK. [8]Yorkshire Regional Genetics Service, Chapel Allerton Hospital, Leeds LS7 4SA, UK. [9]Leicestershire Genetics Centre, University Hospitals of Leicester NHS Trust, Leicester LE1 5WW, UK. [10]Human Genetics, Division of Medical Sciences, University of Dundee, Dundee DD1 9SY, UK. [11]NW Thames Regional Genetics Service, Kennedy Galton Centre, London HA1 3UJ, UK. [12]Peninsula Regional Genetics Service, Royal Devon & Exeter Hospital, Exeter EX1 2ED, UK. [13]SW Thames Regional Genetics Service, St George's Hospital, London SW17 0RE, UK. [14]West Midlands Regional Genetics Service, Birmingham Women's Hospital, Birmingham B15 2TG, UK. [15]Sheffield Regional Genetics Service, Sheffield Children's NHS Foundation Trust, Sheffield S10 2TH, UK. [16]West of Scotland Regional Genetics Service, Laboratory Medicine, Southern General Hospital, Glasgow G51 4TF, UK. [17]South Western Regional Genetics Service, University Hospitals of Bristol NHS Foundation Trust, Bristol BS2 8EG, UK. [18]Northern Genetics Service, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle NE1 3BZ, UK. [19]Faculty of Medicine, University of Southampton, Southampton University Hospitals NHS Trust, Southampton SO16 5YA, UK. [20]Genetic Medicine, Manchester Academic Health Science Centre, St Mary's Hospital, Manchester M13 9WL, UK. [21]Merseyside and Cheshire Clinical Genetics Service, Liverpool Women's NHS Foundation Trust, Liverpool L8 7SS, UK. [22]SE Thames Regional Genetics Service, Guy's and St Thomas NHS Foundation Trust, London SE1 9RT, UK. [23]NE Thames Regional Genetics Service, Great Ormond St Hospital, London WC1N 3JH, UK. [24]University Department of Medical Genetics & Regional Genetics Service, St Mary's Hospital, Manchester M13 9WL, UK. [25]University of Aberdeen and North of Scotland Clinical Genetics Service, Aberdeen Royal Infirmary, Aberdeen AB25 2ZA, UK. [26]Northern Ireland Regional Genetics Service, Belfast HSC Trust, Department of Medical Genetics, Queen's University Belfast, Belfast BT9 7AB, UK. [27]East Anglian Regional Genetics Service, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK. [28]South East of Scotland Clinical Genetics Service, Western General Hospital, Edinburgh EH4 2XU, UK. [29]All Wales Medical Genetics Service, University Hospital of Wales, Cardiff CF14 4XW, UK. [30]Department of Cancer Genetics, Royal Marsden NHS Foundation Trust, Sutton SM2 5PT, UK. [31]Oxford Regional Genetics Service, Oxford University Hospitals NHS Trust, Oxford OX3 7LJ, UK. [32]Department of Gynaecologic Oncology, Royal Marsden NHS Foundation Trust, London SW3 6JJ, UK. [33]University of Queensland Diamantina Institute, University of Queensland, Princess Alexandra Hospital, Woolloongabba, Brisbane 4102, Australia. [34]Clinical Pharmacology and Barts and The London Genome Centre, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London EC1M 6BQ, UK. [35]Oxford Centre for Diabetes, Endocrinology and Medicine, University of Oxford, Churchill Hospital, Oxford OX3 7LJ, UK. [36]Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford OX3 7LJ, UK. [37]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK. [38]Department of Statistics, University of Oxford, Oxford OX1 3TG, UK.
†Lists of participants and their affiliations appear in the Supplementary Information.
*These authors contributed equally to this work.

Fig. 1). To maximize time, sample and cost efficiency we used a pooled approach combining 200 ng of DNA from each of the 24 individuals into a single pool that we hybridized to a custom pull-down containing the DNA repair panel (Supplementary Table 2). We performed sequencing using an Illumina HiSeq2000 sequencer that generated a minimum coverage per pool of 480× for at least 90% of the target region (Supplementary Fig. 2). Sequence variants were called using Syzygy[11], the performance of which was evaluated using previously generated data in a subset of the samples. The sensitivity of base substitution calling was 99.6% (439 out of 439 common variants and 24 out of 26 rare variants that were present in 1 out of 24 individuals in a pool). The sensitivity of insertion/deletion calling was 94.4% (51 out of 54 rare insertions/deletions present in 1 out of 24 individuals in a pool; Supplementary Table 3).

We next considered the 34,564 sequence variants called by Syzygy. We first focused on PTVs because of the strong association of this class of mutation with disease. In total, 1,044 PTVs were called by Syzygy and we used a PTV prioritization method to stratify the genes according to the number of different, rare truncating mutations present within the samples[12]. PPM1D showed the strongest signal in this analysis, and we confirmed by Sanger sequencing that five individuals carried different PPM1D PTVs. Two of these individuals had ovarian cancer in addition to breast cancer.

To explore the role of PPM1D in breast and ovarian cancer susceptibility further, we next performed a case-control Sanger sequencing analysis of PPM1D in a total of 13,642 individuals—7,781 unrelated

individuals with breast and/or ovarian cancer and 5,861 population controls (Supplementary Table 1). We initially sequenced all PPM1D exons and intron–exon boundaries, but after completing this analysis in 3,803 samples we noted that all 10 PTV mutations identified occurred within the last exon of PPM1D, and this clustering was highly significant ($P = 8.2 \times 10^{-6}$). We thus analysed the remaining 9,839 samples for this mutation cluster region, identifying a further 16 PTVs (Supplementary Table 1 and Fig. 1). In total, we identified 25 PPM1D PTVs in individuals with breast and/or ovarian cancer and 1 in controls ($P = 1.12 \times 10^{-5}$; Figs 1 and 2a and Supplementary Table 4). This included 18 mutations in 6,912 individuals with breast cancer ($P = 2.42 \times 10^{-4}$) and 12 mutations in 1,121 individuals with ovarian cancer ($P = 3.10 \times 10^{-9}$). The histological features of the cancers in PPM1D-mutation carriers were diverse, and five individuals had both breast and ovarian cancer (Supplementary Table 5). The case series included 773 individuals with mutations in BRCA1 or BRCA2 (termed BRCA1/2-mutation carriers), four of whom also carried PTVs in PPM1D (4 out 773 versus 1 out of 5,861 controls, $P = 8.30 \times 10^{-4}$). We also identified a total of 16 non-synonymous, 14 synonymous and 1 intronic variant across the cases and controls; there was no evidence for an association with cancer for these variant classes (Supplementary Table 6).

The Sanger sequencing chromatograms for the PPM1D PTVs were unusual for heterozygous mutations as the mutant allele was considerably and consistently lower than the wild-type allele, suggesting the mutations were mosaic in lymphocyte DNA (Fig. 2a and Supplementary Fig. 3). This contrasted with the non-truncating variants
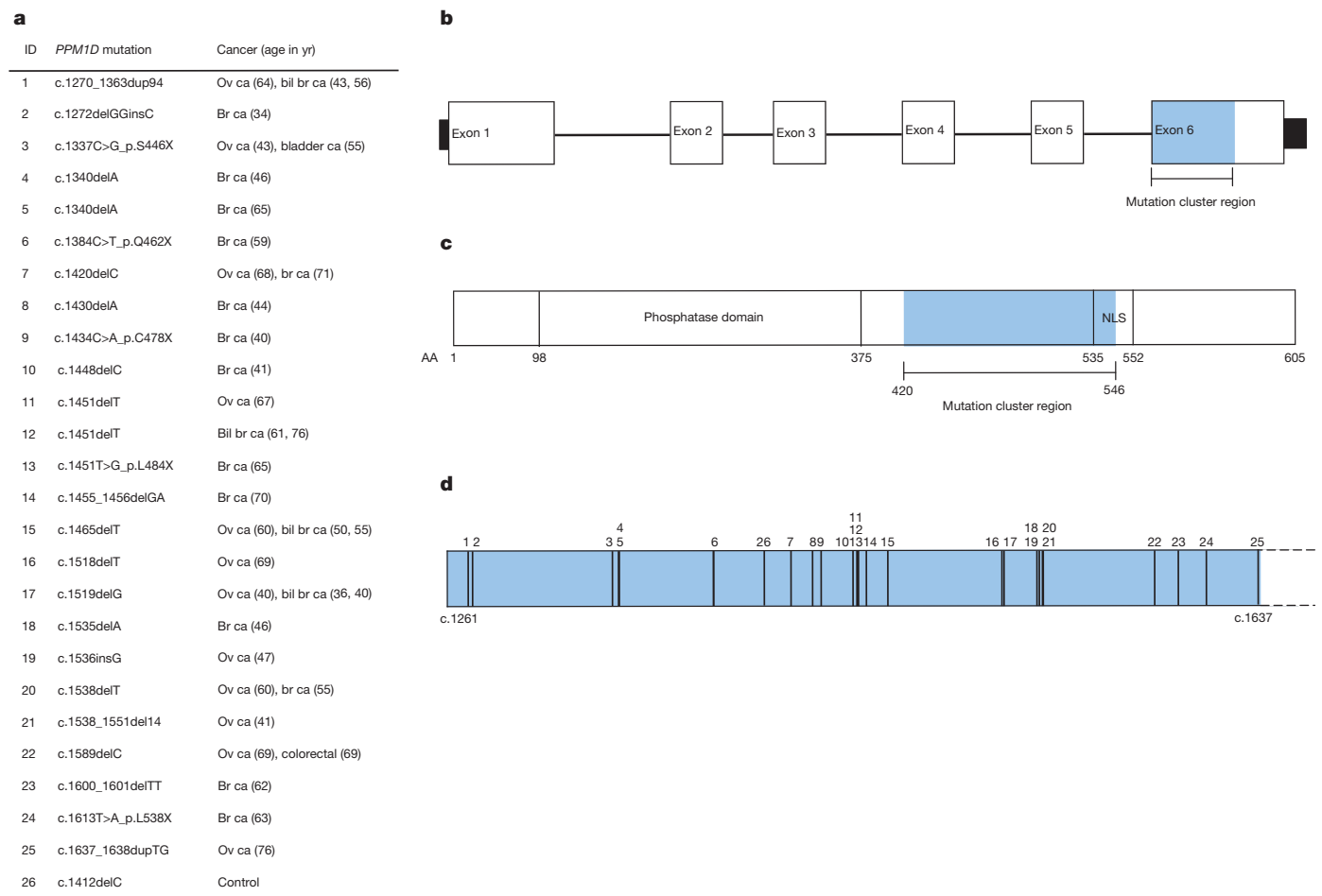

a

ID	PPM1D mutation	Cancer (age in yr)
1	c.1270_1363dup94	Ov ca (64), bil br ca (43, 56)
2	c.1272delGGinsC	Br ca (34)
3	c.1337C>G_p.S446X	Ov ca (43), bladder ca (55)
4	c.1340delA	Br ca (46)
5	c.1340delA	Br ca (65)
6	c.1384C>T_p.Q462X	Br ca (59)
7	c.1420delC	Ov ca (68), br ca (71)
8	c.1430delA	Br ca (44)
9	c.1434C>A_p.C478X	Br ca (40)
10	c.1448delC	Br ca (41)
11	c.1451delT	Ov ca (67)
12	c.1451delT	Bil br ca (61, 76)
13	c.1451T>G_p.L484X	Br ca (65)
14	c.1455_1456delGA	Br ca (70)
15	c.1465delT	Ov ca (60), bil br ca (50, 55)
16	c.1518delT	Ov ca (69)
17	c.1519delG	Ov ca (40), bil br ca (36, 40)
18	c.1535delA	Br ca (46)
19	c.1536insG	Ov ca (47)
20	c.1538delT	Ov ca (60), br ca (55)
21	c.1538_1551del14	Ov ca (41)
22	c.1589delC	Ov ca (69), colorectal (69)
23	c.1600_1601delTT	Br ca (62)
24	c.1613T>A_p.L538X	Br ca (63)
25	c.1637_1638dupTG	Ov ca (76)
26	c.1412delC	Control

Figure 1 | Clustering of cancer predisposing mutations in PPM1D.
a, PPM1D mutations and cancer phenotype. Bil br ca, bilateral breast cancer; br ca, breast cancer; ov ca, ovarian cancer. Age at cancer diagnosis is indicated in parentheses. **b**, PPM1D gene with the region targeted by mutations (mutation cluster region) in blue. **c**, PPM1D protein showing position of mutation cluster region downstream of the phosphatase domain and upstream/overlapping the nuclear localization signal (NLS). **d**, PPM1D mutation cluster region showing the position of mutations. The numbers above give the position of the mutations and correspond to the IDs in **a**.
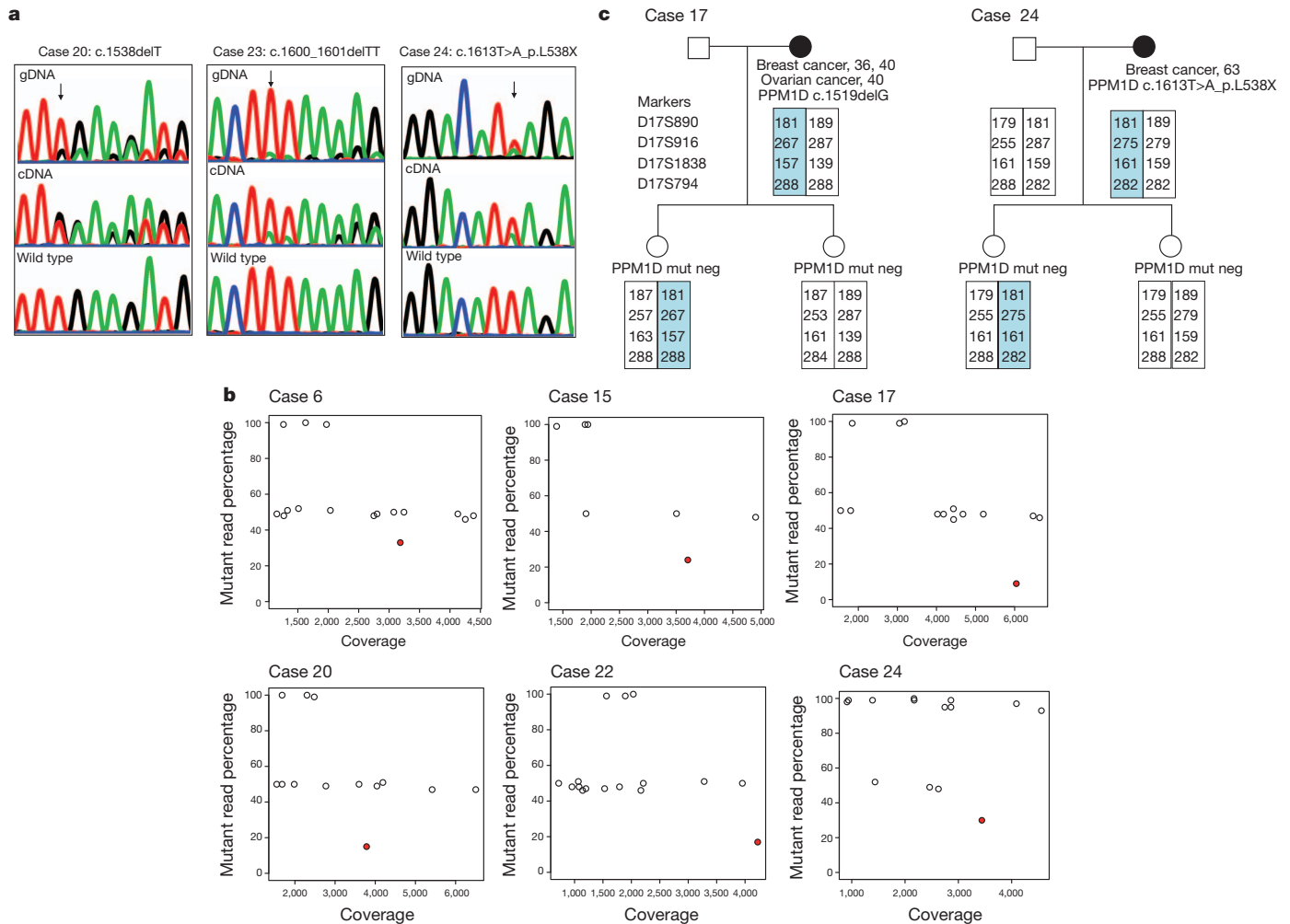
Figure 2 | *PPM1D* mutations are mosaic in lymphocyte DNA. a, Sanger sequencing traces showing mutant allele is lower in genomic DNA (gDNA) extracted from peripheral blood lymphocytes than is typical for heterozygous mutations. The cDNA analysis demonstrates that the mutations lead to truncated products, not nonsense-mediated RNA decay. **b,** Deep PCR amplicon sequencing showing heterozygous *BRCA1/2* variants at 50% (open dots), whereas the *PPM1D* mutation is present at a lower percentage (red dots). **c,** Haplotype analysis in two families. The offspring of *PPM1D*-mutation carriers have different maternal haplotypes spanning *PPM1D* (highlighted), but neither of them carries the mutation, indicating that it is either not present, or mosaic in the germline of the proband.

that all had normal sequencing profiles. DNA from saliva was available for two individuals and the PTVs were present at similar amplitude to that identified in the corresponding blood-derived DNA (Supplementary Fig. 3). To confirm that the PTV mutations were bona fide we used two other mutation detection methods: deep PCR amplicon sequencing[13] (Fig. 2b, Supplementary Fig. 4 and Supplementary Table 4) and multiplex ligation-dependent probe amplification (MLPA)[14] (Supplementary Fig. 5 and Supplementary Table 7). For the deep PCR amplicon sequencing we generated Nextera libraries of pooled PCR products covering *BRCA1*, *BRCA2* and the *PPM1D* mutation, which we sequenced using an Illumina MiSeq sequencer generating a median coverage of 3,387× across the *PPM1D* mutation (Supplementary Fig. 4 and Supplementary Table 4). This confirmed that the *PPM1D* PTVs were present at a lower proportion than heterozygous polymorphisms in *BRCA1* and *BRCA2*, with a median mutant read percentage of 16% (range 5–34%). Furthermore, we sequenced the original DNA repair panel in six cases individually (that is, unpooled), which again confirmed the mutations were present, but mosaic (Supplementary Fig. 6 and Supplementary Table 4). For three samples we had data from both the deep PCR amplicon sequencing and the DNA repair panel that gave identical mutation percentage results (Supplementary Table 4). Finally, family studies were also consistent with mosaicism; none of the 14 relatives carried the *PPM1D* mutation identified in the proband. Most compellingly, for each of probands 17 and 24, we identified two

offspring that had inherited different maternal haplotypes at the *PPM1D* locus, but neither offspring carried the relevant maternal *PPM1D* mutation, demonstrating that the mutations were either not present or mosaic in the germline of the probands (Fig. 2c).

PPM1D (protein phosphatase, Mg²⁺/Mn²⁺ dependent 1D; also known as wild-type p53-induced phosphatase 1 or *Wip1*) was first identified in a screen for p53 target genes induced by ionizing radiation[15]. *PPM1D* encodes a 605-amino-acid protein with an amino-terminal phosphatase catalytic domain and a C-terminal domain that contains a putative nuclear localization signal[16] (Fig. 1). *PPM1D* transcription is upregulated in response to various types of DNA damage in a p53-dependent manner. Once upregulated, PPM1D has been shown to dephosphorylate and downregulate several targets, particularly proteins associated with the DNA damage response (DDR) pathway initiated by the ATM (ataxia telangiectasia mutated) or ATR (ATM- and Rad3-related) protein kinases, and including tumour suppressors with a proven role in cancer susceptibility such as p53 (ref. 17), ATM[18] and CHK2 (ref. 19). Thus, it has been proposed that a primary role of PPM1D is as a homeostatic regulator of the DDR pathway, facilitating return of cells to their normal state after repair of damaged DNA[17]. There is also accumulating evidence that *PPM1D* is involved in onco-genesis[16]. *PPM1D* amplification and overexpression have been demonstrated in several human tumours[20], including breast cancers[21] and ovarian clear cell carcinoma[22], and is a promising therapeutic target[22–24].

The clustering of PTVs within the 370-base-pair (bp) region corresponding to amino acids 420–546, which is downstream of the phosphatase catalytic domain but precedes or disrupts the nuclear localization signal[25], suggests the PTVs are not acting as simple loss-of-function mutations (Fig. 1). Moreover, all the PTVs were in the last exon and thus predicted to evade nonsense-mediated RNA decay and result in a truncated protein that retains the phosphatase catalytic domain, rather than in haploinsufficiency[25,26]. We confirmed this experimentally for three mutations (Fig. 2a). To investigate the effect of PPM1D PTVs we generated complementary DNA expression constructs representing two mutant alleles (PPM1D c.1384C>T, case 6, and PPM1D c.1420delC, case 7) and tested their ability to suppress p53 activation in response to ionizing radiation exposure. As expected, the normal increase in p53 levels after exposure to ionizing radiation was moderately suppressed in human U2OS tumour cells transfected with a wild-type PPM1D expression construct, matching previous observations[16,17] (Fig. 3). The suppression of p53 was enhanced in cells transfected with the mutant PPM1D expression constructs, suggesting that each of these alleles encodes a hyperactive PPM1D isoform, that is, consistent with a gain-of-function rather than loss-of-function effect (Fig. 3). Similar effects were also observed in HeLa and 293 cells (Supplementary Fig. 7).

To investigate the mechanism of oncogenesis in PPM1D PTV-mutation carriers we analysed eight tumours from five individuals. Intriguingly, the PPM1D mutations were not detectable in any of the tumours by Sanger sequencing or MLPA (Supplementary Fig. 8). Through microsatellite analysis we confirmed that the tumours were from the correct individuals and demonstrated loss of heterozygosity at the PPM1D locus in seven out of eight tumours, although there was no evidence of PPM1D copy number alteration (Supplementary Fig. 8 and Supplementary Table 8). We microdissected stromal tissue from the ovarian tumour in four cases and undertook deep sequencing across the PPM1D PTV in blood, tumour and stromal DNA. Each mutation was present in the blood, at similar levels to that detected previously, absent from the tumour and either absent (two cases) or present at very low levels (5 out of 915 and 4 out of 5,793 reads) in the stroma, consistent with lymphocyte contamination (Supplementary Fig. 8 and Supplementary Table 5).

These data are intriguing and strongly suggest the mechanism of cancer association in PPM1D-mutation carriers differs from that in carriers of mutations in other DNA repair genes associated with predisposition to these cancers. There are several potential explanations. It is possible the mutation was present in the cell of cancer origin but was subsequently lost, perhaps because a PPM1D mutation can act as a driver to initiate oncogenesis, but is not required, or is detrimental to the progression of the resulting cancer. The allele loss we observed at the PPM1D locus could be interpreted as supportive of this hypothesis, but it should be noted that it is not known if the lost allele carried the PPM1D PTV, and loss in this region of 17q is common in these cancers. Alternatively, the absence of the PPM1D mutation in the tumour could be because oncogenesis is being driven by the mutation in circulating blood cells. Another possibility is that the PPM1D mutations are not directly involved in causing breast or ovarian cancer. For example, they could be a separate manifestation of an underlying lesion, perhaps one that causes genomic instability, which can lead to selection and clonal expansion of cells with PPM1D PTVs and also to cancers in other tissues. Clearly, further studies will be required to explain the mechanism of oncogenesis in PPM1D-mutation carriers.

Irrespective of the mechanism of the association, our data demonstrate that individuals with mosaic PPM1D PTVs in the mutation cluster region are at increased risk of cancer. The association is not explicable by increasing age, unlike recently reported mosaic chromosomal abnormalities[27,28] (Supplementary Table 5). To estimate the cancer risks we undertook a retrospective cohort analysis using information on breast and ovarian cancer occurrence in the 6,577 unrelated individuals negative for BRCA1/2 mutations and controls, by modelling the retrospective likelihood of the observed mutation
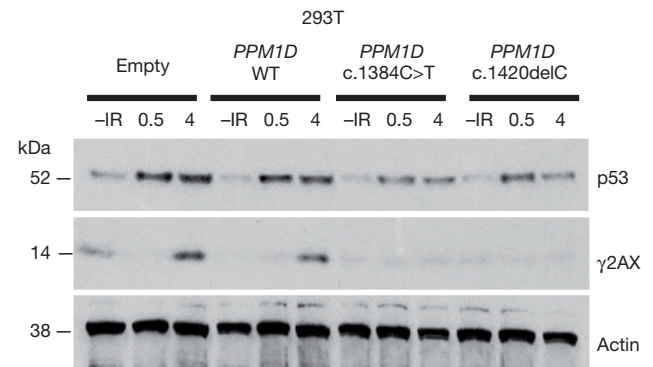

Figure 3 | The effect of mutant PPM1D isoforms on p53 activation. p53 wild-type U2OS human osteosarcoma cells were transfected with PPM1D cDNA expression constructs and exposed to ionizing radiation (IR; 5 Gy). At 30 min and 4 h intervals whole-cell lysates were analysed by western blot to estimate the ionizing-radiation-induced activation of p53. Western blots showing p53 and actin (loading control) protein levels at different times (in hours) after ionizing radiation exposure are shown. γH2AX is a marker of DNA damage, including double-stranded DNA breaks. 'Empty' denotes transfection with empty construct. PPM1D WT denotes transfection with the wild-type PPM1D construct. PPM1D c.1384C>T and PPM1D c.1420delC denote transfection with mutant PPM1D constructs. Suppression of p53 was enhanced in cells transfected with mutant constructs, suggesting that these alleles encode hyperactive PPM1D isoforms.

status conditional on the disease phenotype, as previously described[9,29]. This approach adjusts for our ascertainment of cases with more extreme phenotypes such as young age of onset or bilateral breast cancer, which we use to empower gene discovery[4–7,10,30]. The relative risk of breast cancer for PPM1D PTV-mutation carriers was estimated to be 2.7 (95% confidence interval: 1.3–5.3; $P = 5.38 \times 10^{-3}$), which translates to approximately 23% cumulative risk by age 80. The relative risk of ovarian cancer was estimated to be 11.5 (95% confidence interval: 4.3–30.4; $P = 9.95 \times 10^{-7}$), which translates to approximately 18% cumulative risk by age 80. It is noteworthy that we sequenced an unselected hospital-based series of 322 patients with ovarian cancer in whom we identified five PPM1D PTVs, suggesting that 1–2% of patients with ovarian cancer may contain mosaic PPM1D mutations.

The frequency of PPM1D PTVs in BRCA1/2-mutation carriers with breast and/or ovarian cancer was also significantly different from population controls (4 out of 773 versus 1 out of 5,861; $P = 8.30 \times 10^{-4}$) and similar to that in cases of breast and/or ovarian cancer without BRCA1/2 mutations (4 out of 773 versus 21 out of 6,634; $P = 0.56$), suggesting that PPM1D PTVs are also associated with increased risks of cancer in BRCA1/2-mutation carriers. Studies of unselected, population-based patients with cancer and of larger series of BRCA1/2-mutation carriers would be of value to extend our observations, and to explore further the prevalence and cancer risks associated with PPM1D mutations.

These data provide new insights into ovarian and breast cancer, potentially identifying a new class of genetic defect that lies somewhere between classic germline genetic predisposition mutations and tumour-specific somatic events. It is also highly plausible that PPM1D mutations are associated with other cancers, and broad evaluation of individuals with other tumour types would be of interest. More generally, the clinical implications of a mosaic cancer predisposition marker that is genetic, but not hereditary, and that is detectable in the blood but not the tumour(s) it is associated with are rather profound, particularly if this phenomenon is observed in other genes/contexts.

Our results also provide insights into genetic variation, particularly in relation to the nature and effect of rare gene mutations associated with disease. Given the truncating mutations we report probably have a gain-of-function effect, the widespread interchangeable use of the terms 'truncating mutation' and 'loss-of-function mutation' is inappropriate. We believe a more descriptive term such as 'protein-truncating variant', which does not imply the functional consequence of the mutation, is preferable. We also provide evidence that mosaic mutations can

have relevance to common disease. Such variants are challenging to detect by Sanger sequencing, but are detectable by next-generation sequencing approaches. It is therefore likely that further examples of mosaic disease-associated mutations will be forthcoming, although studies to define the frequency and characteristics of mosaic mutations in control individuals will be essential, to ensure the implications of such variants in case series are correctly interpreted. Finally, although newer sequencing technologies are making large-scale whole-genome sequencing experiments ever more feasible, it is likely that focused sequencing experiments with tailored design and analytical prioritization strategies, such as those used here, will have utility over the next few years.

METHODS SUMMARY

Lymphocyte DNA from 8,046 individuals affected with breast and/or ovarian cancer and 5,861 population-based controls were included. A custom pull-down that included 507 genes (DNA repair panel) was designed using the Agilent SureSelect target enrichment system and sequenced in samples from 1,150 women, in pools of 24 samples, with an Illumina HiSeq2000. Sequence reads were mapped to the human reference genome (hg19) using BWA (version 0.5.6). Variant calling was undertaken with Syzygy[11] (version 1.2.4). Primers for *PPM1D* Sanger sequencing were designed using exon-primer. Amplicons were sequenced using the BigDye terminator cycle sequencing kit and an ABI3730 automated sequencer (ABI PerkinElmer). Deep PCR amplicon sequencing was undertaken by amplifying the *PPM1D* mutation cluster region, and *BRCA1* and *BRCA2* using a Multiplex PCR kit (Qiagen), preparing libraries with Nextera technology[11] and sequencing on an Illumina MiSeq. Sequencing of the DNA repair panel in indexed samples from six *PPM1D* PTV-mutation carriers was undertaken using Illumina TruSeq kits for library preparation and an Illumina HiSeq2000. For these latter experiments, sequence reads were mapped using Stampy version 1.0.14 and variants were called with Platypus (http://www.well.ox.ac.uk/platypus). MLPA was undertaken using the SALSA MLPA probe mix P200 (MRC Holland). Microsatellite analysis was undertaken with 5'6-FAM tagged primer pairs. PCR products were run on a 3730xL genetic analyser (ABI PerkinElmer) and data were analysed using GeneMarker v1.51 (SoftGenetics). The U2OS, HeLa and HEK293 (p53 wild-type) cell lines were transfected with a plasmid containing full-length wild-type *PPM1D* cDNA and the *PPM1D* open reading frame subcloned into pCMV6-AN-HA (Origene), generating a construct that could express a PPM1D N-terminal haemagglutinin (HA) epitope fusion protein. Mutations were introduced using the QuickChange II XL site-directed mutagenesis kit (Stratagene). Statistical analyses were performed using the statistics package in R. Cancer risks were estimated within a retrospective cohort analysis framework[9,29].

Full Methods and any associated references are available in the online version of the paper.

1. Cirulli, E. T. & Goldstein, D. B. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature Rev. Genet.* **11,** 415–425 (2010).
2. Turnbull, C. & Rahman, N. Genetic predisposition to breast cancer: past, present, and future. *Annu. Rev. Genomics Hum. Genet.* **9,** 321–345 (2008).
3. Gayther, S. A. & Pharoah, P. D. The inherited genetics of ovarian and endometrial cancer. *Curr. Opin. Genet. Dev.* **20,** 231–238 (2010).
4. Rahman, N. *et al.* PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nature Genet.* **39,** 165–167 (2007).
5. Renwick, A. *et al.* ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nature Genet.* **38,** 873–875 (2006).
6. Seal, S. *et al.* Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nature Genet.* **38,** 1239–1241 (2006).
7. Meijers-Heijboer, H. *et al.* Low-penetrance susceptibility to breast cancer due to CHEK2*1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nature Genet.* **31,** 55–59 (2002).
8. Meindl, A. *et al.* Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nature Genet.* **42,** 410–414 (2010).
9. Loveday, C. *et al.* Germline RAD51C mutations confer susceptibility to ovarian cancer. *Nature Genet.* **44,** 475–476 (2012).
10. Loveday, C. *et al.* Germline mutations in RAD51D confer susceptibility to ovarian cancer. *Nature Genet.* **43,** 879–882 (2011).
11. Rivas, M. A. *et al.* Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nature Genet.* **43,** 1066–1073 (2011).
12. Snape, K. *et al.* Predisposition gene identification in common cancers by exome sequencing: insights from familial breast cancer. *Breast Cancer Res. Treat.* **134,** 429–433 (2012).
13. Caruccio, N. Preparation of next-generation sequencing libraries using Nextera technology: simultaneous DNA fragmentation and adaptor tagging by *in vitro* transposition. *Methods Mol. Biol.* **733,** 241–255 (2011).
14. Schouten, J. P. *et al.* Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res.* **30,** e57 (2002).
15. Fiscella, M. *et al.* Wip1, a novel human protein phosphatase that is induced in response to ionizing radiation in a p53-dependent manner. *Proc. Natl Acad. Sci. USA* **94,** 6048–6053 (1997).
16. Lu, X. *et al.* The type 2C phosphatase Wip1: an oncogenic regulator of tumor suppressor and DNA damage response pathways. *Cancer Metastasis Rev.* **27,** 123–135 (2008).
17. Lu, X., Nguyen, T. A. & Donehower, L. A. Reversal of the ATM/ATR-mediated DNA damage response by the oncogenic phosphatase PPM1D. *Cell Cycle* **4,** 4060–4064 (2005).
18. Shreeram, S. *et al.* Regulation of ATM/p53-dependent suppression of myc-induced lymphomas by Wip1 phosphatase. *J. Exp. Med.* **203,** 2793–2799 (2006).
19. Fujimoto, H. *et al.* Regulation of the antioncogenic Chk2 kinase by the oncogenic Wip1 phosphatase. *Cell Death Differ.* **13,** 1170–1180 (2006).
20. Bulavin, D. V. *et al.* Amplification of PPM1D in human tumors abrogates p53 tumor-suppressor activity. *Nature Genet.* **31,** 210–215 (2002).
21. Natrajan, R. *et al.* Tiling path genomic profiling of grade 3 invasive ductal breast cancers. *Clin. Cancer Res.* **15,** 2711–2722 (2009).
22. Tan, D. S. *et al.* PPM1D is a potential therapeutic target in ovarian clear cell carcinomas. *Clin. Cancer Res.* **15,** 2269–2280 (2009).
23. Rayter, S. *et al.* A chemical inhibitor of PPM1D that selectively kills cells overexpressing PPM1D. *Oncogene* **27,** 1036–1044 (2008).
24. Hayashi, R. *et al.* Optimization of a cyclic peptide inhibitor of Ser/Thr phosphatase PPM1D (Wip1). *Biochemistry* **50,** 4537–4549 (2011).
25. Chuman, Y. *et al.* PPM1D430, a novel alternative splicing variant of the human PPM1D, can dephosphorylate p53 and exhibits specific tissue expression. *J. Biochem.* **145,** 1–12 (2009).
26. Silva, A. L., Ribeiro, P., Inacio, A., Liebhaber, S. A. & Romao, L. Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay. *RNA* **14,** 563–576 (2008).
27. Jacobs, K. B. *et al.* Detectable clonal mosaicism and its relationship to aging and cancer. *Nature Genet.* **44,** 651–658 (2012).
28. Laurie, C. C. *et al.* Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nature Genet.* **44,** 642–650 (2012).
29. Barnes, D. R. *et al.* Evaluation of association methods for analysing modifiers of disease risk in carriers of high risk mutations. *Genet. Epidemiol.* **36,** 274–291 (2012).
30. Antoniou, A. C. & Easton, D. F. Polygenic inheritance of breast cancer: implications for design of association studies. *Genet. Epidemiol.* **25,** 190–202 (2003).

Supplementary Information is available in the online version of the paper.

Author Contributions E.R., K.S., P.H., N.M.W., T.-P.Y., M.A.B., M.J.C., C.T., J.T., M.I.M., P.de., P.Do. and N.R. (chair) are the Wellcome Trust Case Control Consortium (WTCCC) exon-resequencing group who devised and funded phase 1. J.W.A., J.Ba., J.Be., A.F.B., C.B., G.B., C.C., J.C., R.D., A.D., F.D., D.E., D.G.E., L.G., A.H., L.I., A.K., F.L., Z.M., P.J.M., J.P., M.P., M.T.R., S.Sh., L.W. and N.R. are centre leads of the Breast and Ovarian Cancer Susceptibility Collaboration (BOCS), which is coordinated by M.W.-P. and A.Z. A full list of the WTCCC and BOCS consortia is in the Supplementary Information. R.H. and M.G. assembled the unselected ovarian cancer series. L.G. coordinated phase 1 sequencing. P.H., M.R. and P.Do. undertook analysis of the pooled DNA repair panel. J.D., A.M., S.Se., S.H. and E.R. undertook NGS sequencing and analysis. S.Se. E.R., S.D.V.S., N.A.P., A.Re., K.S., C.L. and J.D. undertook Sanger sequencing, MLPA and tumour microsatellite analysis. E.R. undertook sample selection and data analyses with C.T. A.C.A. wrote the risk analysis software and oversaw the penetrance analysis that was performed by E.R. A.Ri. provided and optimized Platypus. D.N.R., A.C.-F. and J.S.R.-F. undertook histopathological analyses and performed microdissections. I.B., R.B., C.J.L. and A.A. undertook functional analyses. E.R., K.S. and N.R. managed and oversaw all aspects of the study and wrote the manuscript.

METHODS

Patients and samples. We used lymphocyte DNA from 8,046 individuals affected with breast and/or ovarian cancer that were recruited by two studies. There were 7,724 cases recruited through 24 genetics centres in the United Kingdom via the Breast and Ovarian Cancer Susceptibility (BOCS) study, which recruits women ≥18 years who have had breast cancer and/or ovarian cancer and have a family history of breast cancer and/or ovarian cancer. Each proband was screened for *BRCA1* and *BRCA2* mutations (by Sanger sequencing and/or hetero-duplex analysis) and large rearrangements (by MLPA). The remaining 322 cases are an unselected hospital-based series of women with ovarian cancer who were recruited during treatment for ovarian cancer at the Royal Marsden Hospital. The DNA was extracted from peripheral blood samples except in 11 cases, for which DNA was extracted from a lymphoblastoid cell line (note that all of the *PPM1D* mutations were identified in peripheral blood-derived DNA). At least 97% of families were of European ancestry, that is, comparable to the controls. Informed consent was obtained from all participants. The research was approved by the London Multicentre Research Ethics Committee (MREC/01/2/18).

For the phase 1 pooled DNA repair panel experiment we used lymphocyte DNA from 1,150 women with breast cancer, of which 69 also had ovarian cancer. Seventy-eight of these individuals had one mutation, and one individual had two mutations, in known cancer predisposition genes. These were included as positive controls to evaluate variant calling (see below). For the *PPM1D* case-control sequencing experiment we used 7,781 individuals with breast and/or ovarian cancer. We did not use the case data from the pooled DNA repair panel experiment in the case-control analysis, first because the mutation status of individuals cannot be definitively obtained from the pooled experiment as one cannot be certain that every sample is equally represented in a pool, and second because the mutation detection method was different to that used in the case-control experiment. We used our standard case and control sample trays for the case-control *PPM1D* sequencing experiment and the sample selection was blind to the pooled DNA repair panel experiment. There were 885 individuals that were part of both experiments.

Samples and pathology information from mutation-positive families. For families in which a *PPM1D* mutation was detected, we sought DNA samples from relatives. We also requested tumour material, histopathology information and immunohistochemical profiles, including hormone receptor and HER2 status for patients with breast cancer, in probands from the hospitals where they had been treated. Representative tumour blocks were retrieved where possible and examined by two histopathologists, and classified and graded according to the World Health Organization 2003 classification[31,32]. Tumours were microdissected under a stereomicroscope and genomic DNA was extracted from tumours and, where possible, stroma using the DNeasy kit (Qiagen) as previously described[32].

Controls. We used lymphocyte DNA from 5,861 population-based controls obtained from the 1958 Birth Cohort Collection, a continuing follow-up of persons born in the United Kingdom in one week in 1958. Biomedical assessment was undertaken during 2002–2004 at which blood samples and informed consent were obtained for creation of a genetic resource but phenotype data for these individuals is not available. At least 97% of the controls were of European ancestry (http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=724&sitesectiontitle=Welcome+to+the+1958+National+Child+Development+Study).

DNA repair panel sequencing. We identified genes for inclusion on the DNA repair panel from the Gene Ontology (GO) database (http://www.geneontology.org/) using the search term 'DNA repair' (GO accession 0006281) and from STRING (http://string-db.org/) by identifying all genes interacting with *ATM*, *BRCA1*, *BRCA2*, *BRIP1*, *CHEK2* and *PALB2* with highest confidence (≥0.9). This data set was manually curated to remove duplicate genes and pseudogenes. Consensus CoDing Sequence (CCDS) transcripts for the remaining genes were retrieved from UCSC Genome Browser (http://genome.ucsc.edu/, from November 2010) (Supplementary Table 2). Genomic coordinates for all coding exons were identified and targeted in a custom pull-down designed using the Agilent SureSelect target enrichment system[33]. We created 48 pools of DNA that each included 4 μl of 50 ng μl^{-1} of DNA (that is, 200 ng) from 24 individuals. We sheared 80 μl of the pooled DNA using Covaris technology. We prepared libraries without gel size selection or PCR enrichment using the Illumina genomic PE sample prep kit (Ilumina) and performed target enrichment according to the Agilent SureSelect protocol. Sequencing was performed by the Wellcome Trust Centre for Human Genetics high-throughput DNA sequencing and MRC hub in Oxford on an Illumina HiSeq2000 (v2 flow cell, one lane of sequencing per pool) generating 2×100-bp reads. Sequence reads for each pool were mapped to the human reference genome (hg19) using BWA (version 0.5.6)[34]. Mapped reads were filtered to remove ambiguous alignments with a quality score of 0 and bases with a call quality below 22 were masked. Of the remaining reads for each pool, 50–60% fell within the target regions, except for pool 21 in which the on-target percentage was significantly lower.

Median coverage for each pool achieved for target regions after filtering was between $2,849\times$ and $5,545\times$. This corresponded to an average coverage of 119–231$\times$ per sample. All pools had 90% of the target covered at a minimum of 480$\times$. Target regions within the MHC achieved substantially lower coverage and were excluded from further analysis.

We also sequenced the DNA repair panel in six *PPM1D* PTV-positive individuals using Illumina TruSeq kits for library preparation to enable sample indexing. Genomic DNA (1.5 μg) was fragmented and the libraries prepared using the Illumina TruSeq sample preparation kit (index set A). One pool of six libraries (500 ng each) was enriched as before but with the addition of extra blocking primers targeted against the TruSeq index adaptor sequences. Sequencing was performed at the Institute of Cancer Research with an Illumina HiSeq2000 (v3 flow cell, one lane) generating 2×100-bp reads. Mapped reads were filtered to remove ambiguous alignments with a quality score of 0 and bases with a call quality below 22 were masked. Of the remaining reads, 41–43% fell within the target region for each individual. Median coverage of the target for each individual after filtering was between 602$\times$ and 690$\times$. All individuals had 90% of the target covered at a minimum of 50$\times$.

***PPM1D* Sanger sequencing.** We designed primers to PCR-amplify and Sanger-sequence *PPM1D* using exon-primer from the UCSC genome browser (http://genome.ucsc.edu/, from November 2010). Primers and conditions are available on request. PCR reactions were performed using the Qiagen multiplex PCR kit. Amplicons were unidirectionally sequenced using the BigDye terminator cycle sequencing kit and an ABI3730 automated sequencer (ABI PerkinElmer). We analysed the full coding sequence in 2,456 cases and 1,347 controls. As all of the mutations identified in these samples were restricted to exon 6 we sequenced the mutation cluster region (c.1261-20-c.1695), but not the rest of the gene, in the remaining 5,325 cases and 4,514 controls. We also sequenced the mutation cluster region in all available samples from relatives of *PPM1D* PTV-positive probands. All sequencing traces were independently analysed by two individuals who were blinded to the others analysis. Each individual analysed the sequencing with both automated software (Mutation Surveyor, SoftGenetics) and manual visual inspection. All putative mutations were confirmed by bidirectional sequencing from a fresh aliquot of the stock DNA. We also undertook Sanger sequencing of the *PPM1D* cluster region, in triplicate, in DNA from eight tumour samples and four ovarian stromal samples.

For the cDNA sequencing we established lymphoblastoid cell lines from three individuals with *PPM1D* PTVs (cases 20, 23 and 24). RNA was extracted using the RNeasy minikit (Qiagen) and cDNA synthesized using the ThermoScript RT–PCR system (Invitrogen), using standard protocols. We amplified the mutation cluster region using a cDNA-specific primer (forward, 5′-ACCACCAGTCAAGTCACTGG-3′, reverse, 5′-TCTTTCGCTGTGAGGTTGTG-3′), which we sequenced as described above.

Deep PCR amplicon sequencing. In lymphocyte DNA we amplified the *PPM1D* mutation cluster region and the full coding sequence and intron–exon boundaries of *BRCA1* and *BRCA2* using the multiplex PCR kit (Qiagen). We prepared indexed libraries of the PCR products using Nextera technology (Ilumina)[13]. We created two pools of 24 indexed libraries that we sequenced using an Illumina MiSeq, generating 2×150-bp reads. Data from 20 individuals passed quality control coverage metrics, generating median coverage greater than 500$\times$ across the *PPM1D* cluster region (average median coverage 3,384$\times$).

For the tumour analyses we amplified the mutation cluster region in tumour, stroma and blood DNA using an Illumina Nextera XT library preparation kit and supplied protocol (Illumina). To attain the required 1 ng input for tagmentation we also amplified *BRCA1* in 24 samples as described above and we then created one pool of 24 indexed libraries that we sequenced using an Illumina MiSeq, generating 2×150-bp reads. We visually inspected all sequencing reads present at the mutation site after alignment with Stampy to determine whether the *PPM1D* mutation was present.

DNA repair panel data. For the pooled DNA repair panel analysis, variant calling was undertaken with Syzygy (version 1.2.4)[11]. We successfully identified 402 out of 439 previously validated single nucleotide polymorphisms (SNP) with a minor allele frequency (MAF) > 5% genotyped through a breast cancer genome-wide association study (GWAS)[35] with high confidence, and the remaining 37 SNPs were detected at lower confidence. Syzygy also detected 75 out of 80 rare variants (MAF < 1%) included in the study as positive controls (24 out of 26 base substitutions, 14 out of 14 insertions, 30 out of 32 deletions, and 7 out of 8 complex indels; Supplementary Table 3). Thus, sensitivity was 99.6% for base substitutions and 94.4% for rare indels. Frequency estimation for rare variants was assessed by evaluation of 39 *BRCA1* and *BRCA2* variants at a frequency of one per pool. Syzygy correctly estimated the frequency in 33 of the 35 variants it detected, incorrectly estimating the frequency at two per pool for the remaining two variants.

Deep PCR amplicon sequencing data. For the deep PCR amplicon sequencing and the indexed DNA repair panel sequencing in six individuals, sequence reads were

mapped to the human reference genome (hg19) using Stampy version 1.0.14 (ref. 36). Duplicate reads were flagged using Picard version 1.60 (http://picard.sourceforge.net). Variant calling was performed with Platypus version 0.1.9 (http://www.well.ox.ac.uk/platypus)[37]. The mutant read percentage was calculated as the proportion of total reads at the variant location that contained the variant, with a minimum mutant read percentage threshold of 5%.

Variant annotation. Annotation for all experiments was undertaken with reference to CCDS transcripts from Ensembl version 65 identified using a custom Perl script (Supplementary Table 2). Variant calls were annotated for changes with respect to the chosen transcript and assigned a consequence type from the list used by Ensembl.

PTV prioritization method. This is a gene-based (rather than the more typical variant-based) strategy that aims to prioritize potential disease-associated genes for follow-up by leveraging two properties of protein truncating variants: (1) the strong association of rare truncating variants with disease, and (2) collapsibility. Different PTVs within a gene typically result in the same functional effect and can be combined equally. We implemented the method using the statistics package in R. We first outputted all the predicted protein truncating variants: stop gains, coding frameshifts and essential splice site variants (-2, -1, $+1$, $+2$ and $+5$). For this experiment we defined 'rare' as PTVs that were seen only once in the DNA repair panel data. We next stratified the genes according to the number of different, rare singleton PTVs called. We excluded genes for which samples had been included as positive controls (Supplementary Table 3). *PPM1D* was the top gene in this analysis. We are undertaking further analyses and follow-up of the DNA repair panel data, which we aspire to present in a separate publication.

MLPA. We designed 22 probe pairs targeting *PPM1D* PTVs ($n = 18$), wild-type *PPM1D* ($n = 2$), wild-type *BRCA1* ($n = 1$) and wild-type *CEP112* ($n = 1$) (Supplementary Table 7). We added the synthetic probes to the SALSA MLPA probe mix P200 (MRC Holland). MLPA reactions were performed in triplicate according to the manufacturer's instructions. MLPA was undertaken in lymphocyte DNA from 17 probands and in eight tumour DNA samples (from five individuals). In brief, probes were hybridized to 150 ng of denatured DNA, amplified by PCR, and separated on an ABI 3130 genetic analyzer (Applied Biosystems). Data were analysed using GeneMarker v1.51 software (SoftGenetics).

Microsatellite analysis. We used 5′6-FAM tagged primer pairs and PCR conditions for 17q microsatellite analysis as listed in Supplementary Table 8. Ten microlitres of a mastermix of 30 μl ROX size standard and 1 ml HiDi formamide were added to each reaction post PCR, denatured at 95 °C for 5 min, and cooled to −20 °C for 5 min. Reactions were run on a 3730xL genetic analyser (Applied Biosystems) under the fragment analysis protocol. Data were analysed using GeneMarker v1.51 software (SoftGenetics). Microsatellite analysis was undertaken in lymphocyte DNA from 13 individuals from eight families, and in eight tumour DNA samples and four stroma DNA samples from five individuals. Of note, one of these cases (case 17) contained both *BRCA1* and *PPM1D* mutations. Both genes are located at chromosome 17q and it is the wild-type *BRCA1* allele that is reduced in the tumours and therefore the relevance of the loss of heterozygosity with respect to *PPM1D* is difficult to deduce.

Cell line and plasmid constructs. The U2OS, HeLa and HEK293 (all p53 wild-type) cell lines were obtained from the American Type Culture Collection (ATCC). Cells were cultured and maintained according to the supplier's instructions. Cells were transfected with plasmid DNA using Lipofectamine 2000 (Invitrogen). A plasmid containing full-length wild-type *PPM1D* cDNA (pCMV6 entry-PPM1D) was obtained from Origene, and the *PPM1D* open reading frame (ORF) subcloned into pCMV6-AN-HA (Origene), generating a construct that could express a PPM1D N-terminal haemagglutinin (HA) epitope fusion protein. Truncating mutations were introduced into the *PPM1D* ORF of this construct using the QuickChange II XL site-directed mutagenesis kit (Stratagene). The following DNA amplication primers were used: *PPM1D* mutant 1 (c.1384C>T): forward, 5′-AGAGAATGTCTAAGGTGTAGTC-3′, reverse, 5′-GACTACACCTTAGACA TTCTCTC-3′; *PPM1D* mutant 2 (c.1420delC): forward, 5′-GATCCAGAACCA TTGAAG-3′, reverse, 5′-CTTCAATGGTTCTGGATC-3′.

Western blot analysis of p53 levels. U2OS, HeLa and HEK293 cells were transfected with *PPM1D* expression constructs, and 24 h after transfection, cells were exposed to γ-irradiation (5 Gy) from an X-ray source. Whole cell lysates were generated from transfected cells after irradiation (at 30 min and 4 h time points) and subjected to protein electrophoresis. Immunoblotting of electrophoresed lysates was performed using antibodies specific for p53 (9282S, Cell Signaling Technology) and actin (sc-1616, Santa Cruz Biotech).

Frequency and risk estimation. Statistical analyses were performed using the statistics package in R. The significance of mutation clustering was modelled under a binomial distribution in which the probability of observing a mutation in the last

exon, which comprises 31% of the coding sequence, was 0.31. The frequency in *BRCA1/BRCA2*-mutation carriers and non-carriers was compared using a two-sided test of proportions. Risk estimation was implemented using a competing risks retrospective likelihood model incorporating age at onset according to a proportional hazards model. Because individuals screened for *PPM1D* mutations were selected on the basis of both personal and family history of breast or ovarian cancer, standard methods of analysis that ignore the sampling frame would yield biased estimates of the risk ratios. To address this, we analysed data within a retrospective cohort approach by modelling the conditional likelihood of the observed genotypes given the disease phenotypes, using information on breast and ovarian cancer occurrence in the set of 6,577 unrelated individuals negative for *BRCA1/2* mutations (*BRCA1/2* mutation-positive individuals from the BOCS series and all the unselected ovarian case series were excluded) and controls. Male controls were included in the analysis, but were not considered to be at risk of developing breast or ovarian cancer. We assumed a competing risks model, under which, each individual was at risk of developing breast or ovarian cancer. This has been shown to provide unbiased estimates of the risk ratios for breast and ovarian cancer in which a genetic variant may be associated with one or both of the diseases[29]. We estimated the *PPM1D*-mutation carrier frequency in the population and breast and ovarian cancer risk ratios simultaneously. Because mutation screened probands may have been selected on the basis of bilateral breast cancer diagnosis or on the basis of both breast and ovarian cancer diagnosis we allowed for the risks of breast or ovarian cancer diagnosis after the first cancer diagnosis, including the risk of contralateral breast cancer. This model assumes that the increased breast cancer (including contralateral) or ovarian cancer risk after the first cancer diagnosis is entirely due to the susceptibility as defined by the model, with no other variation in risk. Site-specific cancer risks were assumed to be independent conditional on genotype. Therefore, the incidence of cancer at the second site was assumed to be the same as if the preceding cancer had not occurred, with the exception of contralateral breast cancer incidence after the first breast cancer, which was assumed to be half the overall breast cancer incidence, as only one breast was at risk. In all models females were censored at age 80 years. We assumed that the breast and ovarian cancer incidences depend on the underlying *PPM1D* genotype through models of the form: $\lambda(t) = \lambda_0(t) \exp(\beta x)$, in which $\lambda_0(t)$ is the baseline incidence at age t in non-mutation carriers, β is the log-risk ratio associated with the mutation, and x takes value 0 for non-mutation carriers and 1 for mutation carriers. The overall breast and ovarian cancer incidences, over all genotypes, were constrained to agree with the population incidences for England and Wales in the period of 1993–1997 (ref. 38), as described previously[30,39]. The models were parameterized in terms of the mutation frequencies and log-risk ratios for breast and ovarian cancer. Parameters were estimated using maximum likelihood estimation and were implemented in the pedigree analysis software MENDEL[40]. The variances of the parameters were obtained by inverting the observed information matrix. To obtain confidence intervals for the risk ratios and perform hypothesis testing, log-risk ratios were assumed to be normally distributed. A Wald test-statistic was used to test the null hypothesis that $\beta = 0$ for both breast and ovarian cancer. Because *PPM1D* mutations were not found to segregate within families, we did not take into account precise family histories or pedigree information and therefore did not incorporate the effects of other susceptibility genes.

31. Tavassoli, F. A. & Devilee, P. in *Pathology and Genetics of Tumours of the Breast and Female Genital Organs* (IARC, 2003).
32. Hernandez, L. *et al.* Genomic and mutational profiling of ductal carcinomas in situ and matched adjacent invasive breast cancers reveals intra-tumour genetic heterogeneity and clonal selection. *J. Pathol.* **227,** 42–52 (2012).
33. Gnirke, A. *et al.* Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature Biotechnol.* **27,** 182–189 (2009).
34. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).
35. Turnbull, C. *et al.* Genome-wide association study identifies five new breast cancer susceptibility loci. *Nature Genet.* **42,** 504–507 (2010).
36. Lunter, G. & Goodson, M. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res.* **21,** 936–939 (2011).
37. Rimmer, A., Mathieson, I., Lunter, G. & McVean, G. Platypus: An Integrated Variant Caller (http://www.well.ox.ac.uk/platypus) (2012).
38. Curado, M. P. *et al.* Cancer incidence in five continents, Volume VIII. *IARC Sci. Publ.* **No 160,** 1–781 (2002).
39. Antoniou, A. C. *et al.* Evidence for further breast cancer susceptibility genes in addition to *BRCA1* and *BRCA2* in a population-based study. *Genet. Epidemiol.* **21,** 1–18 (2001).
40. Lange, K., Weeks, D. & Boehnke, M. Programs for Pedigree Analysis: MENDEL, FISHER, and dGENE. *Genet. Epidemiol.* **5,** 471–472 (1988).

LETTER

Ultrahard nanotwinned cubic boron nitride

Yongjun Tian[1], Bo Xu[1], Dongli Yu[1], Yanming Ma[2], Yanbin Wang[3], Yingbing Jiang[4], Wentao Hu[1], Chengchun Tang[5], Yufei Gao[1], Kun Luo[1], Zhisheng Zhao[1], Li-Min Wang[1], Bin Wen[1], Julong He[1] & Zhongyuan Liu[1]

Cubic boron nitride (cBN) is a well known superhard material that has a wide range of industrial applications. Nanostructuring of cBN is an effective way to improve its hardness by virtue of the Hall–Petch effect—the tendency for hardness to increase with decreasing grain size[1,2]. Polycrystalline cBN materials are often synthesized by using the martensitic transformation of a graphite-like BN precursor, in which high pressures and temperatures lead to puckering of the BN layers[3]. Such approaches have led to synthetic polycrystalline cBN having grain sizes as small as ∼14 nm (refs 1, 2, 4, 5). Here we report the formation of cBN with a nanostructure dominated by fine twin domains of average thickness ∼3.8 nm. This nanotwinned cBN was synthesized from specially prepared BN precursor nanoparticles possessing onion-like nested structures with intrinsically puckered BN layers and numerous stacking faults. The resulting nanotwinned cBN bulk samples are optically transparent with a striking combination of physical properties: an extremely high Vickers hardness (exceeding 100 GPa, the optimal hardness of synthetic diamond), a high oxidization temperature (∼1,294 °C) and a large fracture toughness (>12 MPa m$^{1/2}$, well beyond the toughness of commercial cemented tungsten carbide, ∼10 MPa m$^{1/2}$). We show that hardening of cBN is continuous with decreasing twin thickness down to the smallest sizes investigated, contrasting with the expected reverse Hall–Petch effect below a critical grain size or the twin thickness of ∼10–15 nm found in metals and alloys.

Polycrystalline cubic boron nitride possesses high hardness and high chemical stability and is the best-known material for cutting ferrous and carbide-forming hard substances where diamond completely fails. However, commercial polycrystalline cBNs have grain sizes in the micrometre scale and a Vickers hardness (H_V) of only 33–45 GPa (refs 6, 7), much lower than that of diamond (∼100 GPa)[8]. The application of cBN is therefore limited. For decades, scientists have designed various ways to increase the hardness of polycrystalline cBN[1,2,4,5] while retaining its excellent chemical stability.

The Hall–Petch effect[9,10], through which material hardness increases with decreasing grain size, has been widely observed. For polycrystalline cBN or diamond, reducing the grain size has been the most widely used approach for enhancing hardness. Conditions of high pressure and high temperature (HPHT) are required for the synthesis because temperature enhances the nucleation rate, and pressure suppresses grain growth by reducing the atomic diffusion responsible for crystallization[2,11]. Experimentally, nanograined (ng) diamonds with grain sizes of 10–30 nm have been synthesized[11]. High Knoop hardnesses of up to 110–140 GPa have been reported. Efforts in reducing the grain size further to below 10 nm have been unsuccessful as a result of intergranular fractures in these bulk ng-diamonds[12]. For cBN, the best attempts[1,2] performed so far have produced ng-cBN samples at pressures ∼20 GPa with a minimum grain size of 14 nm and a record hardness of 85 GPa. This grain size is close to the critical value at which the Hall–Petch relation breaks down, as demonstrated in metals and alloys[13–16]. It remains a technical challenge to synthesize polycrystalline cBNs with perfect grain boundaries and smaller grain sizes, because

numerous high-energy grain boundaries in a polycrystalline material generate large driving forces for grain growth.

Fine twin domains can be stabilized within materials capable of twinning. Coherent twin boundaries possess excess energy that is typically about one order of magnitude lower than that of grain boundaries[15]. This makes nanotwinned structures energetically more stable than their nanograined counterparts, thereby allowing finer control of domain size. Twin boundaries have been experimentally verified to have a hardening effect identical to that of high-angle grain boundaries[15], because, similarly to grain boundaries, twin boundaries also serve as barriers against gliding dislocations. Hence, the synthesis of nanotwinned substructures within grains may be a feasible route to achieving ultrafine microstructural sizes approaching a few nanometres. In practice, however, a strategically designed experimental protocol essential to fine-tune nanotwin substructures has yet to be achieved.

The use of graphite-like BN as precursors merely generates ng-cBN[1,2,4,5], although twin domains are occasionally observed during the growth of cBN single crystals as undesired defects[17,18]. Here we report the synthesis of ubiquitously nanotwinned cBN (nt-cBN) with an average twin thickness of 3.8 nm under HPHT conditions, using a special turbostratic onion-like BN (oBN) nanoparticle precursor material.

The starting oBN nanoparticles (∼30–150 nm in diameter) consisted of concentric BN spherical shells (Fig. 1a, b) with abundant puckering and stacking faults. The inter-shell spacing was determined by X-ray diffraction at 0.351 nm (Supplementary Fig. 2a), which is slightly larger than the inter-layer spacings in hexagonal (0.3328 nm) and rhombohedral (0.334 nm) BN[17]. The oBN precursors were pressed into a pellet (2.5 mm in diameter and 1 mm in height) and subjected to HPHT treatments with a multi-anvil apparatus. X-ray diffraction and Raman spectroscopy were performed on the recovered samples to characterize structural phase transformations (Supplementary Fig. 2). At 15 GPa, no transformation was observed at temperatures below 1,000 °C. Transformation into a translucent cubic phase mixed with a small amount of the wurtzite phase occurred above 1,200 °C. The sample transformed entirely into transparent cBN (Fig. 1c and Supplementary Fig. 1) above 1,600 °C. Other synthetic trials performed at different pressures from 12 GPa to 25 GPa showed a similar temperature dependence of the phase transformation. For comparison, a considerably higher temperature (2,200 °C) was used to acquire pure cBN when graphite-like BN precursors were employed[1]. The use of the oBN precursor is responsible for the significantly reduced synthetic temperature (by about 600 °C) and the stabilization of nanotwinned microstructure (see below).

The synthetic nt-cBN bulk samples had a diameter of about 2 mm and possessed a unique homogeneous microstructure. Typical transmission electron microscopy (TEM) and high-resolution TEM (HRTEM) images of an as-synthesized bulk sample are shown in Fig. 1d–f. Grains were about the same size as the original onions, irregular in shape but approximately equiaxed. Selected area electron diffraction (inset of Fig. 1d) confirmed that the sample was pure cBN containing randomly oriented nanograins. HRTEM revealed a high

[1]State Key Laboratory of Metastable Materials Science and Technology, Yanshan University, Qinhuangdao 066004, China. [2]State Key Laboratory for Superhard Materials, Jilin University, Changchun 130012, China. [3]Center for Advanced Radiation Sources, University of Chicago, Chicago, Illinois 60439, USA. [4]TEM Laboratory, University of New Mexico, Albuquerque, New Mexico 87131, USA. [5]School of Material Science and Engineering, Hebei University of Technology, Tianjin 300130, China.
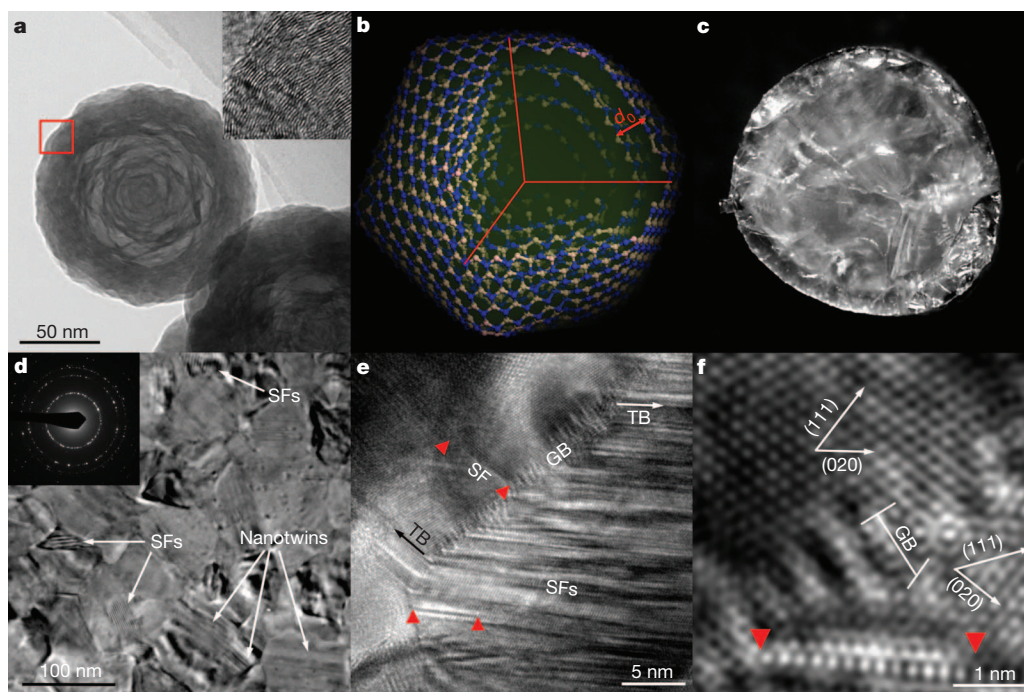
Figure 1 | Starting oBN nanoparticles and nt-cBN bulk synthesized at 15 GPa and 1,800 °C. a, TEM image of oBN nanoparticles. Inset, HRTEM image corresponding to the position marked with the red box, showing defects including lattice puckering, bending and stacking faults. **b**, Schematic icosahedral model of a five-shell oBN nanoparticle. d_0 is the inter-shell spacing. **c**, Photograph of an nt-cBN bulk sample with a diameter of about 2 mm. **d**, TEM image of a typical microstructure in nt-cBN. Nanotwins and stacking faults (SFs) are marked. Inset, a selected area electron diffraction pattern. **e**, HRTEM image of nt-cBN showing Shockley partial dislocations (red triangles) emitted from grain boundaries (GB) and a high density of stacking faults in twin domains. TB, twin boundary. **f**, Enlarged HRTEM image of **e**, showing the orientation relationship between adjacent nanograins.

density of dislocations at grain boundaries, as well as a high density of stacking faults within the twin domains. Narrow and sharp high-angle grain boundaries about five or six atomic layers thick were frequently present (Fig. 1e, f). All nanograins contained densely spaced lamellar {111} twins (Fig. 2a, b). Figure 2c depicts the twin-thickness distribution derived from 726 counts of 60 grains based on TEM and HRTEM images. Twin thicknesses were predominantly below 10 nm, with an average of 3.8 nm. As a result of the low excess energy of twin boundaries, variations in duration (from several minutes to half an hour) during the HPHT treatments did not noticeably affect the average and distribution of twin thickness, indicating that the nanotwins are highly stable.

The H_V of bulk nt-cBN samples was measured with a standard square-pyramidal diamond indenter. Reliable hardness values are best determined from the asymptotic hardness region through a well-controlled indentation process[19,20]. We recorded the variations in H_V

with respect to a series of applied loads (Fig. 3). The asymptotic hardness obtained at loads above 3 N reached an extremely high value of 108 GPa, which is the highest hardness reported so far for polycrystalline cBN and exceeds even that of synthetic diamond. Overall, the hardness measured for different nt-cBN samples synthesized between 1,800 and 1,950 °C at 12–15 GPa ranged from 95 GPa to 108 GPa. For direct comparison, we measured the hardness of commercial cBN single crystals with the same technique and obtained values lower than 49 GPa with loads of up to 4.9 N, above which the crystals fractured. The results are in good agreement with other reported hardness data (30–43 GPa) on cBN single crystals[21]. Our measured Knoop and nanoindentation hardnesses of nt-cBN samples are 77.7 and 98.5 GPa (Supplementary Table 1), respectively.

Radial cracks formed in bulk nt-cBN at loads of 9.8 and 19.6 N were used to determine the fracture toughness (K_{IC}) with the equation[22] $K_{IC} = 7.42 \times 10^{-2} F/L^{1.5}$, where L and F are the length of the crack and
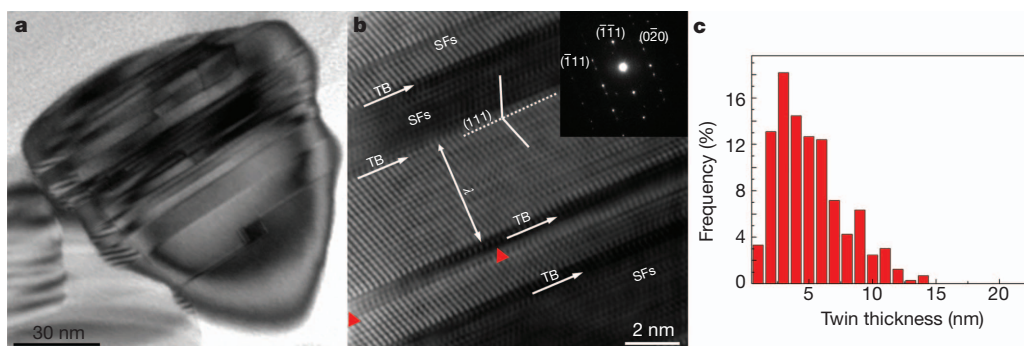
Figure 2 | Microstructure of synthetic nt-cBN. a, Bright-field TEM image of a nanograin. **b**, HRTEM image and corresponding selected area electron diffraction pattern (inset) along the [101] zone axis of the nanograin shown in **a**. Lamellar nanotwins with various thicknesses (λ) are present in the nanograin. The twinning plane is of the {111} type, and the lattice fringe angles across the twin plane are 70.53°. Stacking faults and Shockley dislocations (red triangles) are labelled. **c**, Thickness distribution of the twins measured from TEM and HRTEM images with an average twin thickness of 3.8 nm.
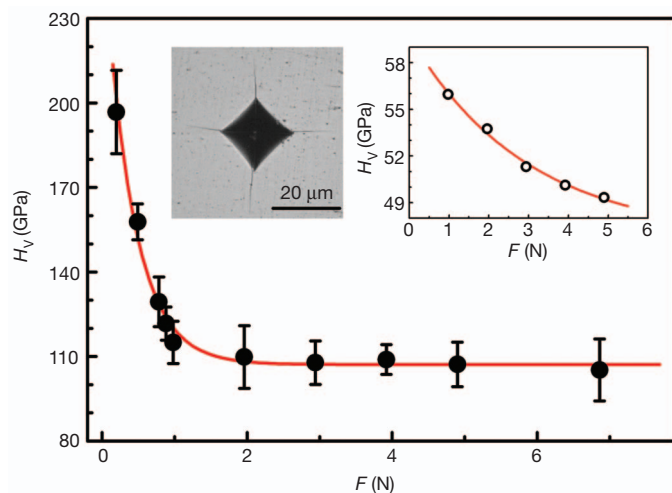
Figure 3 | The H_V of an nt-cBN bulk sample as a function of applied load (F). The H_V of the nt-cBN bulk decreases from ~196 GPa at 0.2 N to its asymptotic value, 108 GPa, beyond 3 N. Error bars indicate s.d. ($n = 5$). Left inset, an optical micrograph of the Vickers indentation with cracks produced at a load of 19.6 N. Right inset, the H_V–F curve of a 0.3-mm cBN single crystal. H_V does not reach its asymptotic value, and the crystal fractures when F exceeds 4.9 N.

Figure 4 | H_V as a function of average grain size (d) or twin thickness (λ) for polycrystalline cBN bulk materials. Experimental data for nt-cBN bulk material (triangle) and for ng-cBN bulk materials[1] (squares) are shown. Using equation (1) to fit the experimental data, we obtained $H_V = 42.6 + 126d^{-1/2} + 130.7d^{-1}$. k (126 GPa nm$^{1/2}$) is taken from a previously reported value[1]. The fitted C (130.7 ± 16.8 GPa nm) parameter characterizing the quantum confinement effect is in excellent agreement with the theoretical value (136 GPa nm) from $211N_e^{1/3}\exp(-1.191f_i)$ for cBN. This coincidence is not accidental and may provide proof of the existence of the quantum confinement effect in the synthesized nt-cBN. Inset, the yield strength as a function of λ for nt-Cu, in which the critical λ is about 15 nm (ref. 15).

the applied load, respectively. The resulting K_{IC} reached a remarkably high value of 12.7 MPa m$^{1/2}$, which is about 4.5-fold that of cBN single crystal (2.8 MPa m$^{1/2}$) and almost twice that of micrometre-grain-sized cBN (6.8 MPa m$^{1/2}$) (ref. 2). This K_{IC} is 21% greater than that of ng-cBN (10.5 MPa m$^{1/2}$) (ref. 2) and 27% greater than the fracture toughness (10 MPa m$^{1/2}$) of commercial cemented tungsten carbide, a material widely used in the industry as a cutting tool[23].

The thermal stability of nt-cBN was characterized by using heat flow curves measured in air by differential scanning calorimetry (Supplementary Fig. 3). An onset oxidation temperature of ~1,294 °C was determined from the exothermic trough, which is higher than those of single-crystal cBN (~1,103 °C), ng-cBN (~1,187 °C) (ref. 2) and commercial polycrystalline cBN (~1,000 °C) (ref. 4). The high thermal stability of nt-cBN samples, apparently due to the nanotwinned structure, allows higher-temperature industrial applications.

Figure 4 shows that the hardness of polycrystalline cBN increases monotonically with decreasing microstructural size (grain size for grains and/or twin thickness for twins). On the basis of our recent theoretical model[20], the hardness dependence of the grain size (d) for cBN obeys

$$H_V = H_{HP} + H_{QC} = H_0 + kd^{-1/2} + Cd^{-1} \qquad (1)$$

where H_{HP} ($= H_0 + kd^{-1/2}$) and H_{QC} ($= Cd^{-1}$) represent dislocation-related hardening based on the Hall–Petch effect[9,10] and bandgap-related hardening based on the speculated quantum confinement effect[24], respectively. H_0 is the single-crystal hardness and k is a material constant. C is a material-specific parameter equal to zero for metals, and equal to $211N_e^{1/3}\exp(-1.191f_i)$ for covalent materials[20], where N_e is the valence electron density and f_i is the Phillips ionicity of the chemical bond[25]. Both Hall–Petch and quantum confinement effects have been used to account for the hardening behaviour in ng-cBN with grain sizes down to 14 nm (ref. 1).

Our present results show that these two hardening mechanisms remain valid when the twin thickness is significantly reduced to 3.8 nm for cBN (Fig. 4). This finding on covalent cBN is in sharp contrast with those in metals, for which both yield strength and hardness decrease significantly (inset of Fig. 4) and the Hall–Petch hardening mechanism becomes invalid[14,15,26] when the twin thickness or grain size decreases to a critical value (typically ~10–15 nm). The continuously hardening behaviour with decreasing microstructural

sizes (down to 3.8 nm) in cBN may be rationalized as follows. For nanotwins with thicknesses below 3.8 nm, the quantum confinement effect (which is inoperative in metals) becomes dominantly large according to equation (1) (see also the caption of Fig. 4). At a twin thickness of 3.8 nm, each twin lamella contains only about ten unit cells of cBN. This featured nanoscale geometry, as well as the very strong covalent B–N bonding, severely confines the migration of twin boundaries, which is known to induce hardness softening in nanotwinned metals[26].

The use of oBN nanoparticles as precursors is the key to synthesizing nt-cBN, and it highlights the important structural effects of spherical atomic arrangement in oBN on the formation of nanotwinned substructures. It has been established that coarse-grained graphite-like BN precursors can transform into grained cBN or wurtzite BN through martensitic transformation by means of puckering or buckling of the planar BN layers[3]. In our synthesis, the oBN precursors that we used contain naturally puckered BN layers (Fig. 1b), which make the transformation into cBN easier. The high concentration of puckered layers and defects (here, stacking faults; Fig. 1a) in oBN precursors provides ideal sites for the nucleation of cBN. Moreover, the specific orientation relations between coherent nucleus cBN and parent phase oBN must be reserved during the reconstructive phase transition[27]. As a result, multiple parallel laminated nanotwins form at the sub-grain level. High-density domain boundaries with ideal boundary structures in nanotwins make nt-cBN the most desirable superhard material yet achieved. This new material possesses the combined advantages of ultrahigh hardness, high thermal stability, and extremely high fracture toughness, providing a wide range of new industrial applications.

Our findings introduce a new strategy and direction in the quest for superhard materials. The synthesis of nanotwinned substructures by selecting suitable starting materials and optimal P–T conditions, and the use of hardening effects from interlocked nanotwins, are essential. On the basis of our results, if nanotwins at similar scales can be reproduced in polycrystalline diamond (isostructural to cBN and known for twinning in large crystals), an ultrahigh hardness exceeding that of ng-diamond would be achievable.

METHODS SUMMARY

Pure oBN nanospheres with diameters about 30–150 nm were synthesized with a chemical vapour deposition technique[28]. A schematic icosahedral model of a five-shell oBN nanoparticle was relaxed from the nested bucky onion of $B_{30}N_{30}$, $B_{120}N_{120}$, $B_{270}N_{270}$, $B_{480}N_{480}$ and $B_{750}N_{750}$ fullerenes. Geometric optimization was performed with the Forcite module of Materials Studio software, and the Universal Force Field was used. HPHT experiments were performed with a 10-MN two-stage large-volume multi-anvil system at Yanshan University, identical to that described in ref. 29. The standard COMPRES 10/5 sample assembly, consisting of a 10-mm spinel ($MgAl_2O_4$) octahedron with a Re heater and a $LaCrO_3$ thermal insulator, was used[30]. Temperature was measured *in situ* with Type C W-Re thermocouples, and pressures were estimated from previously obtained calibration curves at different temperatures[30]. Recovered bulk samples were about 2 mm in diameter and 0.2–0.5 mm in height. Microstructures of the bulks were characterized by TEM (JEOL-2010F) with an accelerating voltage of 200 kV at the University of New Mexico and TEM (JEM-2010) with an accelerating voltage of 200 kV at Yanshan University. Phase identification was performed with an X-ray powder diffractometer (D8 Discover) with Cu K_α radiation, and by Raman spectroscopy (Renishaw inVia). H_V and Knoop hardness (H_K) were measured with a microhardness tester (KB 5 BVZ). At least five hardness data points were taken for each load, and the hardness values were determined from the asymptotic-hardness region. Nanohardness (H_N) and Young's modulus (E) were measured with a Hysitron TI 950 TriboIndentor with Berkovich indenter at a peak load of 50 mN (see Supplementary Fig. 4 for details). Studies on oxidation resistance were performed in air by differential scanning calorimetry (NETZSCH STA 449 C) with a heating rate of $10\,°C\,min^{-1}$ in the temperature range 20–1,500 °C.

1. Dubrovinskaia, N. *et al.* Superhard nanocomposite of dense polymorphs of boron nitride: noncarbon material has reached diamond hardness. *Appl. Phys. Lett.* **90**, 101912 (2007).
2. Solozhenko, V. L., Kurakevych, O. O. & Le Godec, Y. Creation of nanostuctures by extreme conditions: high-pressure synthesis of ultrahard nanocrystalline cubic boron nitride. *Adv. Mater.* **24**, 1540–1544 (2012).
3. Britun, V. F. & Kurdyumov, A. V. Mechanisms of martensitic transformations in boron nitride and conditions of their development. *High Press. Res.* **17**, 101–111 (2000).
4. Sumiya, H., Uesaka, S. & Satoh, S. Mechanical properties of high purity polycrystalline cBN synthesized by direct conversion sintering method. *J. Mater. Sci.* **35**, 1181–1186 (2000).
5. Dub, S. N. & Petrusha, I. A. Mechanical properties of polycrystalline cBN obtained from pyrolytic gBN by direct transformation technique. *High Press. Res.* **26**, 71–77 (2006).
6. Wentorf, R. H., DeVries, R. C. & Bundy, F. P. Sintered superhard materials. *Science* **208**, 873–880 (1980).
7. Liew, W. Y. H., Yuan, S. & Ngoi, B. K. A. Evaluation of machining performance of STAVAX with PCBN tools. *Int. J. Adv. Manuf. Technol.* **23**, 11–19 (2004).
8. Krauss, A. R. *et al.* Ultrananocrystalline diamond thin films for MEMS and moving mechanical assembly devices. *Diamond Related Materials* **10**, 1952–1961 (2001).
9. Hall, E. O. The deformation and ageing of mild steel: III. Discussion of results. *Proc. Phys. Soc. B* **64**, 747–753 (1951).
10. Petch, N. J. The cleavage strength of polycrystals. *J. Iron Steel Inst.* **174**, 25–28 (1953).
11. Irifune, T., Kurio, A., Sakamoto, S., Inoue, T. & Sumiya, H. Ultrahard polycrystalline diamond from graphite. *Nature* **421**, 599–600 (2003).
12. Sumiya, H. & Irifune, T. Hardness and deformation microstructures of nano-polycrystalline diamonds synthesized from various carbons under high pressure and high temperature. *J. Mater. Res.* **22**, 2345–2351 (2007).
13. Schiøtz, J., Di Tolla, F. D. & Jacobsen, K. W. Softening of nanocrystalline metals at very small grain sizes. *Nature* **391**, 561–563 (1998).
14. Schiøtz, J. & Jacobsen, K. W. A maximum in the strength of nanocrystalline copper. *Science* **301**, 1357–1359 (2003).
15. Lu, L., Chen, X., Huang, X. & Lu, K. Revealing the maximum strength in nanotwinned copper. *Science* **323**, 607–610 (2009).
16. Pande, C. S. & Cooper, K. P. Nanomechanics of Hall–Petch relationship in nanocrystalline materials. *Prog. Mater. Sci.* **54**, 689–706 (2009).
17. He, L. L., Akaishi, M. & Horiuchi, S. Structural evolution in boron nitrides during the hexagonal–cubic phase transition under high pressure at high temperature. *Microsc. Res. Tech.* **40**, 243–250 (1998).
18. Oku, T., Hiraga, K., Matsuda, T., Hirai, T. & Hirabayashi, M. Twin structures of rhombohedral and cubic boron nitride prepared by chemical vapor deposition method. *Diamond Related Materials* **12**, 1138–1145 (2003).
19. Brazhkin, V. *et al.* What does 'harder than diamond' mean? *Nature Mater.* **3**, 576–577 (2004).
20. Tian, Y., Xu, B. & Zhao, Z. Microscopic theory of hardness and design of novel superhard crystals. *Int. J. Refract. Met. Hard Mater.* **33**, 93–106 (2012).
21. Harris, T. K., Brookes, E. J. & Taylor, C. J. The effect of temperature on the hardness of polycrystalline cubic boron nitride cutting tool materials. *Int. J. Refract. Met. Hard Mater.* **22**, 105–110 (2004).
22. Solozhenko, V. L., Dub, S. N. & Novikov, N. V. Mechanical properties of cubic BC_2N, a new superhard phase. *Diamond Related Materials* **10**, 2228–2231 (2001).
23. Berger, C., Scheerer, H. & Ellermeier, J. Modern materials for forming and cutting tools—overview. *Materialwiss. Werkstofftech.* **41**, 5–17 (2010).
24. Tse, J. S., Klug, D. D. & Gao, F. M. Hardness of nanocrystalline diamonds. *Phys. Rev. B* **73**, 140102 (2006).
25. Gao, F. M. *et al.* Hardness of covalent crystals. *Phys. Rev. Lett.* **91**, 015502 (2003).
26. Li, X., Wei, Y., Lu, L., Lu, K. & Gao, H. Dislocation nucleation governed softening and maximum strength in nano-twinned metals. *Nature* **464**, 877–880 (2010).
27. Oleinik, G. S., Petrusha, I. A., Danilenko, N. V., Kotko, A. V. & Shevchenko, S. A. Crystal-oriented mechanism of dynamic recrystallization nucleation in cubic boron nitride. *Diamond Related Materials* **7**, 1684–1692 (1998).
28. Tang, C. C., Bando, Y., Huang, Y., Zhi, C. Y. & Golberg, D. Synthetic routes and formation mechanisms of spherical boron nitride nanoparticles. *Adv. Funct. Mater.* **18**, 3653–3661 (2008).
29. Wang, Y. *et al.* The large-volume high-pressure facility at GSECARS: a 'Swiss-army-knife' approach to synchrotron-based experimental studies. *Phys. Earth Planet. Inter.* **174**, 270–281 (2009).
30. Leinenweber, K. D. *et al.* Cell assemblies for reproducible multi-anvil experiments (the COMPRES assemblies). *Am. Mineral.* **97**, 353–368 (2012).

Author Contributions Y.J.T. conceived the project. Y.J.T., B.X., D.L.Y. and Y.B.W. designed the experiments. C.C.T. synthesized oBN precursors, Y.J.T., B.X., D.L.Y., Y.B.W., Y.F.G., K.L. and Z.S.Z. performed the HPHT experiments, Y.B.J. and W.T.H. performed TEM observations, and B.W. performed molecular dynamics simulations. Y.J.T., B.X., D.L.Y., Y.M.M., Y.B.W., L.-M.W., J.L.H. and Z.Y.L. analysed the data. Y.J.T., B.X., Y.M.M. and Y.B.W. co-wrote the paper. Y.J.T., B.X. and D.L.Y. contributed equally to the study. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.J.T. (fhcl@ysu.edu.cn).

LETTER

Pulsed accretion in a variable protostar

James Muzerolle[1], Elise Furlan[2], Kevin Flaherty[3], Zoltan Balog[4] & Robert Gutermuth[5]

Periodic increases in luminosity arising from variable accretion rates have been predicted for some pre-main-sequence close binary stars as they grow from circumbinary disks[1–3]. The phenomenon is known as pulsed accretion and can affect the orbital evolution and mass distribution of young binaries[2,4], as well as the potential for planet formation[5,6]. Accretion variability is a common feature of young stars, with a large range of amplitudes and timescales as measured from multi-epoch observations at optical[7,8] and infrared[9–13] wavelengths. Periodic variations consistent with pulsed accretion have been seen in only a few young binaries via optical accretion tracers[14–16], albeit intermittently with accretion luminosity variations ranging from zero to 50 per cent from orbit to orbit. Here we report that the infrared luminosity of a young protostar (of age about 10^5 years) increases by a factor of ten in roughly one week every 25.34 days. We attribute this to pulsed accretion associated with an unseen binary companion. The strength and regularity of this accretion signal is surprising; it may be related to the very young age of the system, which is a factor of ten younger than the other pulsed accretors previously studied.

We obtained multi-epoch mid-infrared (MIR) observations of the star forming region IC 348 using the Spitzer Space Telescope. Among the roughly 300 pre-main-sequence objects in the cluster, the protostar LRLL 54361 (which we refer to here as L54361) exhibits by far the largest MIR flux variability. We have a total of 81 separate observations of L54361 taken with all three instruments on board Spitzer. The multi-epoch spectral energy distribution (SED) is shown in Fig. 1. The measured bolometric luminosity of the system ranges from about 0.2 to 2.7 solar luminosities ($L_\odot$). The spectral shape remains relatively constant over this range, apart from slightly bluer MIR colours at higher luminosities.

The photometric light curve indicates that the flux variations occur repeatedly throughout the seven-year span of our observations. The two longest contiguous sets of photometry (Fig. 2) reveal a strong pulse signature in which the flux increases by about two magnitudes in as little as a few days, followed by a longer exponential decay over the following few weeks. The combined photometric data set suggests that the variability of L54361 is periodic in nature; the pulse shape revealed by the contiguous warm Spitzer photometry appears in the older data and at other wavelengths, albeit with variations in the pulse width and peak flux. Using several statistical tests, we find that the flux peaks repeat with a robust period of 25.34 ± 0.01 d.

Follow-up multi-epoch imaging taken with the Hubble Space Telescope at near-infrared wavelengths reveals spatially resolved scattered light structures associated with L54361 (Fig. 3). The central source varies with almost the same amplitude and light curve shape at 1.6 μm as seen in the Spitzer data, and the peak occurs exactly as expected given the previously determined period. The geometry of the scattered light is similar to that of other protostars[17], and is probably produced by cavities carved out of an infalling envelope by one or more outflows. The apparent motion of the scattered light indicates a light echo produced as the pulse peak light travels through the outflow cavities, and suggests that the source of the illumination is relatively isotropic.

There are three primary sources of periodicity in young stellar objects: stellar rotation, Keplerian rotation of an inner disk, and the orbital motion of a close binary companion. Stellar rotation can manifest itself via localized hot or cool spots, or via interactions between the stellar magnetic field and the inner disk. In the case of L54361, we reject stellar rotation effects on several grounds: (1) rotation periods for pre-main-sequence stars range from a few days to two weeks[18] (and protostars are typically faster still[19]), all lower than the measured period; (2) dark spots produce sinusoidal light curves, with amplitudes of a few tenths of a magnitude in the optical and declining to longer wavelengths; (3) hot spots tend to produce less obvious periodicity owing to the more stochastic nature of accretion, and can exhibit phase-dependent asymmetric illumination of the circumstellar material as they rotate with the star[20], which we do not see.

Regarding phenomena related to Keplerian rotation of an inner disk, persistent asymmetric structures such as warps in the inner disk can produce periodic obscuration of both single stars[7] and binary systems[21,22]. We argue that the data are not consistent with this scenario in several respects: (1) obscuration localized to the disk plane would not affect light propagating in the perpendicular direction through the outflow cavity; (2) obscuration events produce characteristic light curve 'dips', whereas we see a positive pulse-like shape; (3) as we show below, the MIR and far-infrared flux of L54361 originates
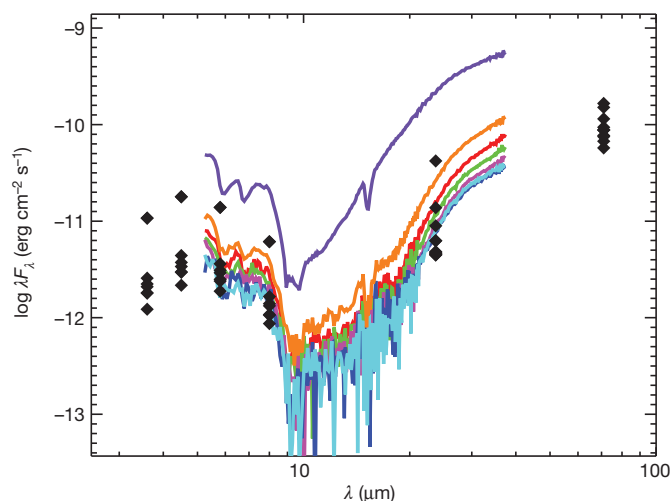

Figure 1 | Multi-epoch spectral energy distribution of L54361. Our complete set of observations taken during cryogenic Spitzer operations are shown, including photometry from all four IRAC channels at 3.6–8 μm and the MIPS 24 and 70 μm channels (diamonds), as well as 7 epochs of IRS spectroscopy (in chronological order: red, green, magenta, blue, cyan, purple and orange lines). Each single-epoch SED exhibits a shape characteristic of class I objects, with the flux rising sharply to longer wavelengths and a strong silicate absorption feature at 8–12 μm. Between epochs, however, the flux varies by as much as an order of magnitude at all wavelengths, with only a slightly shallower spectral slope in the ~15–70-μm continuum as the flux increases.

[1]Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA. [2]National Optical Astronomy Observatory, Tucson, Arizona 85719, USA. [3]Steward Observatory, 933 North Cherry Avenue, University of Arizona, Tucson, Arizona 85721, USA. [4]Max-Planck-Institut fur Astronomie, Konigstuhl 17, 69117 Heidelberg, Germany. [5]Department of Astronomy, University of Massachusetts, Amherst, Massachusetts 01003, USA.

mostly in the infalling envelope, whose total flux would not be significantly affected by localized stellar obscuration.

The third possibility, a connection to binary motion, is plausible in terms of the length of the period of L54361, although we do not yet have direct evidence of a companion. The pulsed accretion scenario could explain both the shape and amplitude of the light curve. Circumbinary disk simulations consistently show gap-clearing by gravitational torques, followed by accretion streams that feed material onto the central stars[1–3]. For certain binary architectures, particularly in the case of a highly eccentric orbit, the stellar accretion depends on orbital phase, with the highest accretion rates typically associated with periastron passages. A qualitatively similar process has also been suggested

for some X-ray binaries, at least one of which has exhibited optical light curves similar to the MIR behaviour of L54361[23,24].

An increase in the accretion luminosity as a result of the binary interaction increases the irradiation heating of circumstellar dust, which then reradiates the energy in the MIR where we observe it. To help test this hypothesis against our observations, we calculated radiative transfer models of protostellar dust emission and scattering[25]. The models include three components that are typical of protostellar systems: infalling envelope, accretion disk and central star. Holding all parameters fixed except for the accretion luminosity, we are able to match the change between SEDs corresponding to two different pulse phases (Fig. 4). The models show a relatively weak wavelength dependence as a function of luminosity, with a slightly flatter spectral slope at about 15 to 70 μm at higher luminosity as a result of optical depth effects in the envelope, in relatively good agreement with the observations.

We do not yet have any direct measure of the central object or its multiplicity status. As our models show, however, the bolometric luminosity provides an estimate of the stellar plus accretion luminosity. Assuming that the low end of the measured range is representative of the stellar luminosity, the combined stellar mass can be roughly estimated by comparing to a theoretical protostellar birthline[26] on an Hertzsprung–Russell diagram. We derive a value of ~0.2 solar masses ($M_\odot$; probably an upper limit because some contribution from accretion is likely, although the luminosity may also be somewhat underestimated because of scattering). Conversely, assuming that the upper end of the range of measured luminosity is due entirely to accretion luminosity and adopting the above stellar mass, we derive a maximum mass accretion rate of $10^{-6}\,M_\odot\,\mathrm{yr}^{-1}$. This is at the upper end of the range of values measured from standard accretion diagnostics[27]. Spectroscopic observations are needed to verify an accretion signature, as well as characterize the binary orbit.

Why L54361 exhibits such a strong and regular signature, unlike the T Tauri-type pulsed accretors observed previously, remains unknown. There may be a connection to its earlier evolutionary stage, in which the infalling envelope provides a steady supply of material to the circumbinary disk. By contrast, T Tauri binaries are older by about a factor of ten, have long since dissipated their natal envelopes, and accrete at lower mean rates. Perhaps stochastic variability from other
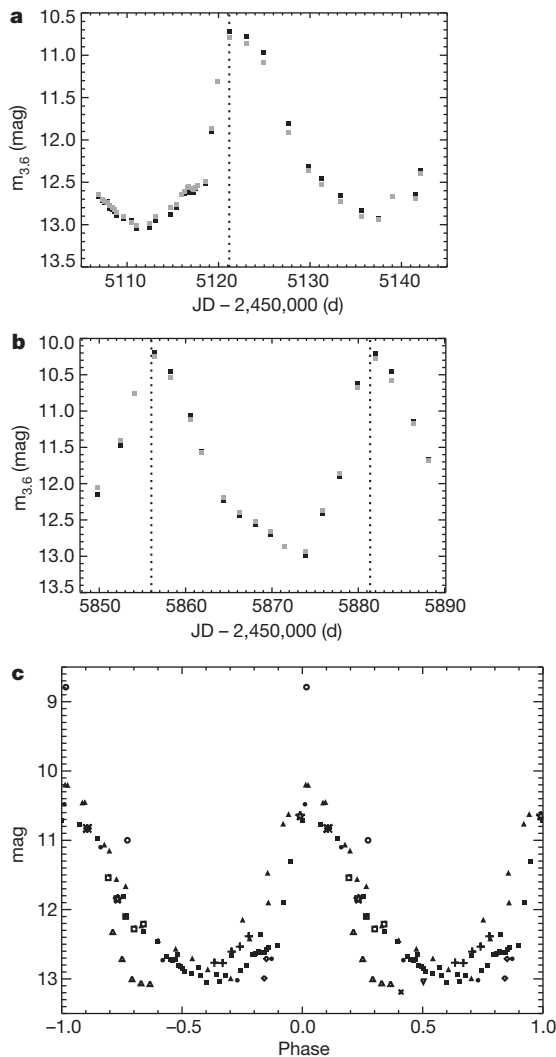

Figure 2 | Photometric light curves for L54361. a, b, IRAC 3.6 μm (black) and 4.5 μm (grey, scaled down to match 3.6) magnitudes from the autumn 2009 (**a**) and autumn 2011 (**b**) Spitzer observing campaigns. Note that the 3σ photometric uncertainties are equal to or smaller than the symbol size. The dashed line in **a** marks the observed peak time, which we set as the fiducial epoch for phase = 0. The dashed lines in **b** mark the predicted peak times assuming a periodicity of 25.34 d. **c**, The phased photometric light curve of L54361, assuming a period of 25.34 d and the phase-zero epoch Julian date (JD) 2,455,121.203. Included are measurements taken at three separate wavelengths. Each symbol type represents a contiguous set of photometry: cryo-Spitzer IRAC 3.6 μm (plus signs, asterisk, cross), warm Spitzer IRAC 3.6 μm (filled squares, filled triangles), MIPS 24 μm (inverted triangle, open diamonds, open stars, open triangles), IRS 24 μm (open circles, open squares) and HST WFC3 1.6 μm (filled circles). The 24 μm photometry values are offset by +7.3 mag and WFC3 photometry offset by −5.3 mag to place everything on the same scale.
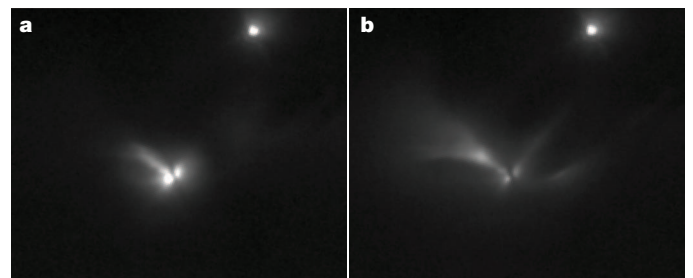

Figure 3 | Near-infrared images of L54361. a, b, Portions of images taken with HST/WFC3 at 1.6 μm at two epochs corresponding to pulse phases of 0 (**a**) and 0.3 (**b**). North is up and east is to the left. L54361 is the extended source just below the centre of the images; the point source at upper right is another young stellar object, LRLL 1843. The light from L54361 subtends roughly 14″ (~4,000 AU at the distance of the IC 348 region) in **a**, and about 50″ (~15,000 AU) in **b**. Most if not all of this light is probably the result of scattering off circumstellar dust in the protostellar envelope. An apparent edge-on disk is visible at the centre of the object, and three separate structures indicative of outflow cavities extend to the northwest, southwest and northeast. The extent and morphology of the scattered light changes substantially between epochs as a result of the propagation of the pulse peak light. (See Supplementary Information for the complete set of HST images.)
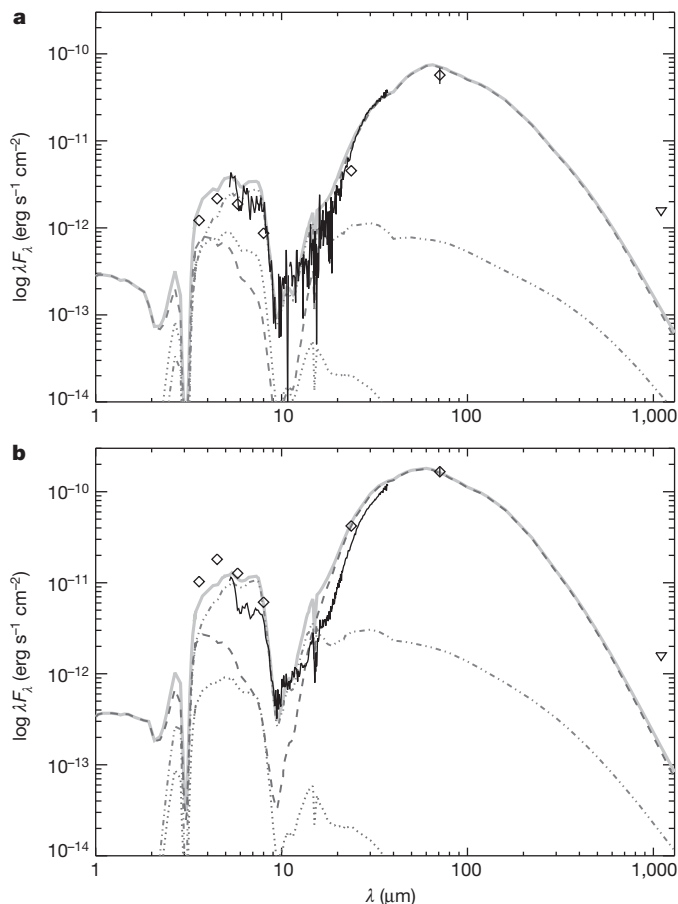
Figure 4 | Protostellar spectral energy distribution models for L54361.
a, **b**, Comparison of models to two sets of observations roughly representing pulse phases of 0.4 (**a**) and 0.15 (**b**). The photometry and spectroscopy are not simultaneous but were selected to correspond roughly to a common flux level. Dashed lines indicate scattering and emission from the infalling envelope; dashed-double-dotted lines represent scattering and emission from the circumstellar accretion disk; dotted lines represent reddened flux from the central star; the solid grey line shows the total flux from all these components. The best-fit parameters include an envelope infall rate of $3 \times 10^{-6} M_\odot \text{ yr}^{-1}$ (assuming a stellar mass of $0.2 M_\odot$), envelope centrifugal radius $R_c = 30$ AU, outflow cavity opening angle $\theta = 30°$, inclination angle of the outflow/stellar rotation axis to the line of sight $i = 70°$, total central luminosity (stellar plus accretion) $L = 0.5 L_\odot$ (**a**) and $1.3 L_\odot$ (**b**), and the fraction of L due to the stellar luminosity $\eta_{star} = 0.5$ (**a**) and 0.19 (**b**). The only parameter that was actually changed between the two models was η_{star}, with appropriate values so that the stellar luminosity remained constant while the accretion luminosity changed by a factor of ~4. (See Supplementary Information for details of the model calculations and parameters.)

sources such as stellar magnetic interactions or disk turbulence can overwhelm the periodic signature in older stars. It is also possible that the particular orbital parameters of L54361 are rare but more favourable for modulating the accretion flow, such as a very large eccentricity.

1. Artymowicz, P. & Lubow, S. H. Dynamics of binary-disk interaction. 1: Resonances and disk gap sizes. *Astrophys. J.* **421**, 651–667 (1994).
2. Artymowicz, P. & Lubow, S. H. Mass flow through gaps in circumbinary disks. *Astrophys. J.* **467**, L77–L80 (1996).
3. Gunther, R. & Kley, W. Circumbinary disk evolution. *Astron. Astrophys.* **387**, 550–559 (2002).
4. Shi, J., Krolik, J. H., Lubow, S. H. & Hawley, J. F. Three dimensional MHD simulation of circumbinary accretion disks: disk structures and angular momentum transport. *Astrophys. J.* **749**, 118–144 (2012).
5. Quintana, E. V. & Lissauer, J. J. Terrestrial planet formation surrounding close binary stars. *Icarus* **185**, 1–20 (2006).
6. Paardekooper, S.-J., Thebault, P. & Mellema, G. Planetesimal and gas dynamics in binaries. *Mon. Not. R. Astron. Soc.* **386**, 973–988 (2008).
7. Bouvier, J. *et al.* Magnetospheric accretion-ejection processes in the classical T Tauri star AA Tauri. *Astron. Astrophys.* **463**, 1017–1028 (2007).
8. Nguyen, D. C. & Scholz, A. van Kerkwijk, M. H., Jayawardhana, R. & Brandeker, A. How variable is accretion in young stars? *Astrophys. J.* **694**, L153–L157 (2009).
9. Flaherty, K. *et al.* Infrared variability of evolved protoplanetary disks: evidence for scale height variations in the inner disk. *Astrophys. J.* **748**, 71–100 (2012).
10. Morales-Calderón, M. *et al.* Ysovar: the first sensitive, wide-area, mid-infrared photometric monitoring of the Orion nebula cluster. *Astrophys. J.* **733**, 50–59 (2011).
11. Liu, M. C. *et al.* Mid-infrared imaging of young stellar objects. *Astrophys. J.* **461**, 334–344 (1996).
12. Barsony, M., Ressler, M. E. & Marsh, K. A. A. Mid-infrared imaging survey of embedded young stellar objects in the ρ Ophiuchi cloud core. *Astrophys. J.* **630**, 381–399 (2005).
13. Kóspál, A. *et al.* The outburst and nature of two young eruptive stars in the North America/Pelican Nebula Complex. *Astron. Astrophys.* **527**, A133 (2011).
14. Basri, G., Johns-Krull, C. M. & Mathieu, R. D. The classical T Tauri spectroscopic binary DQ tau. II. Emission line variations with orbital phase. *Astron. J.* **114**, 781–792 (1997).
15. Mathieu, R. D. *et al.* The classical T Tauri spectroscopic binary DQ tau. I. Orbital elements and light curves. *Astron. J.* **113**, 1841–1854 (1997).
16. Jensen, E. L. N. *et al.* Periodic accretion from a circumbinary disk in the young binary UZ Tau E. *Astron. J.* **134**, 241–251 (2007).
17. Stark, D. P., Whitney, B. A., Stassun, K. & Wood, K. Near-infrared synthetic images of protostellar disks and envelopes. *Astrophys. J.* **649**, 900–913 (2006).
18. Nordhagen, S., Herbst, W., Rhode, K. L. & Williams, E. C. The variability and rotation of pre-main-sequence stars in IC 348: does intracluster environment influence stellar rotation? *Astron. J.* **132**, 1555–1570 (2006).
19. Covey, K. R., Greene, T. P., Doppmann, G. W. & Lada, C. J. The angular momentum content and evolution of class I and flat-spectrum protostars. *Astron. J.* **129**, 2765–2776 (2005).
20. Wood, K. & Whitney, B. Scattered light signatures of magnetic accretion in classical T Tauri stars. *Astrophys. J.* **506**, L43–L45 (1998).
21. Herbst, W. *et al.* The light curve of the weakly accreting T Tauri binary KH 15D from 2005–2010: insights into the nature of its protoplanetary disk. *Astron. J.* **140**, 2025–2035 (2010).
22. Plavchan, P., Gee, A. H., Stapelfeldt, K. & Becker, A. The peculiar periodic YSO WL 4 in ρ Ophiuchius. *Astrophys. J.* **684**, L37–L40 (2008).
23. Skinner, G. K. Observations of optical flares in the recurrent X-ray transient A0538–66. *Nature* **288**, 141–143 (1980).
24. Densham, R. H., Charles, P. A., Menzies, J. W., van der Klis, M. & van Paradijs, J. Four outburst cycles of A0538–66: evidence for a rapidly evolving envelope around the primary. *Mon. Not. R. Astron. Soc.* **205**, 1117–1133 (1983).
25. Kenyon, S. J., Calvet, N. & Hartmann, L. The embedded young stars in the Taurus-Auriga molecular cloud. I — Models for spectral energy distributions. *Astrophys. J.* **414**, 676–694 (1993).
26. Hartmann, L., Cassen, P. & Kenyon, S. J. Disk accretion and the stellar birthline. *Astrophys. J.* **475**, 770–785 (1997).
27. White, R. J. & Hillenbrand, L. A. On the evolutionary status of class I stars and Herbig-Haro energy sources in Taurus-Auriga. *Astrophys. J.* **616**, 998–1032 (2004).

Author Contributions J.M. and K.F. designed the Spitzer observations. Z.B. and R.G. reduced the IRAC images and compiled the photometry, while J.M. reduced and analysed the MIPS data. J.M. and E.F. extracted and analysed the IRS spectroscopy. J.M. designed the HST observations and analysed the images. E.F. calculated the radiative transfer models and fitted the observed SEDs. All authors contributed to the writing of the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.M. (muzerol@stsci.edu).

LETTER

Asymmetric neurotransmitter release enables rapid odour lateralization in *Drosophila*

Quentin Gaudry[1], Elizabeth J. Hong[1], Jamey Kain[2], Benjamin L. de Bivort[2,3] & Rachel I. Wilson[1]

In *Drosophila*, most individual olfactory receptor neurons (ORNs) project bilaterally to both sides of the brain[1,2]. Having bilateral rather than unilateral projections may represent a useful redundancy. However, bilateral ORN projections to the brain should also compromise the ability to lateralize odours. Nevertheless, walking or flying *Drosophila* reportedly turn towards the antenna that is more strongly stimulated by odour[3–5]. Here we show that each ORN spike releases approximately 40% more neurotransmitter from the axon branch ipsilateral to the soma than from the contralateral branch. As a result, when an odour activates the antennae asymmetrically, ipsilateral central neurons begin to spike a few milliseconds before contralateral neurons, and at a 30 to 50% higher rate than contralateral neurons. We show that a walking fly can detect a 5% asymmetry in total ORN input to its left and right antennal lobes, and can turn towards the odour in less time than it requires the fly to complete a stride. These results demonstrate that neurotransmitter release properties can be tuned independently at output synapses formed by a single axon onto two target cells with identical functions and morphologies. Our data also show that small differences in spike timing and spike rate can produce reliable differences in olfactory behaviour.

To navigate towards an odour, an insect can compare the signals from their two antennae and turn in the direction of the stronger signal. Confounding this strategy (by crossing and fixing the antennae, causing them to be spatially reversed) impairs olfactory navigation in bees, ants and locusts[6–8]. *Drosophila* resemble other insects in using this strategy[3–5], but they are unlike other insects in that they have mainly bilateral ORN projections. In the *Drosophila* brain, both ipsilateral and contralateral ORN axons synapse onto antennal lobe projection neurons (PNs)[9], and ipsilateral and contralateral synapses have approximately the same strength[10]. A minority of *Drosophila* olfactory glomeruli receive unilateral ORN projections[1,11], raising the question of whether it is the unilateral glomeruli that enable lateralization.

To investigate lateralization behaviour, we built a spherical treadmill to measure olfactory behaviour in walking *Drosophila* (Fig. 1a and Supplementary Fig. 1). To prevent the head from moving, it was glued to the body, and the two antennae were independently stimulated with odour. When we delivered fermented peach volatiles to one antenna and clean air to the other, flies on the treadmill made a fictive turn towards the odour (Fig. 1b, c). When odour or clean air was delivered symmetrically to the two antennae, flies continued to walk straight.

Fruit volatiles typically activate multiple ORN types[12]. Therefore, we next used a monomolecular odour (pentanoic acid) to target a glomerulus that receives bilateral ORN input more specifically. This odour has just one known high-affinity receptor in the antenna[11,12]. This receptor corresponds to glomerulus DM6, which receives bilateral ORN innervation[2,9]. Pentanoic acid elicited turning behaviour as robust as that elicited by peach volatiles (Fig. 1d). Turning was significantly reduced by a mutation (*Orco*[2]), which silences bilateral ORNs, including the DM6 ORNs (Fig. 1d). The ORN types that project unilaterally do not express the *Orco* gene[11,13],

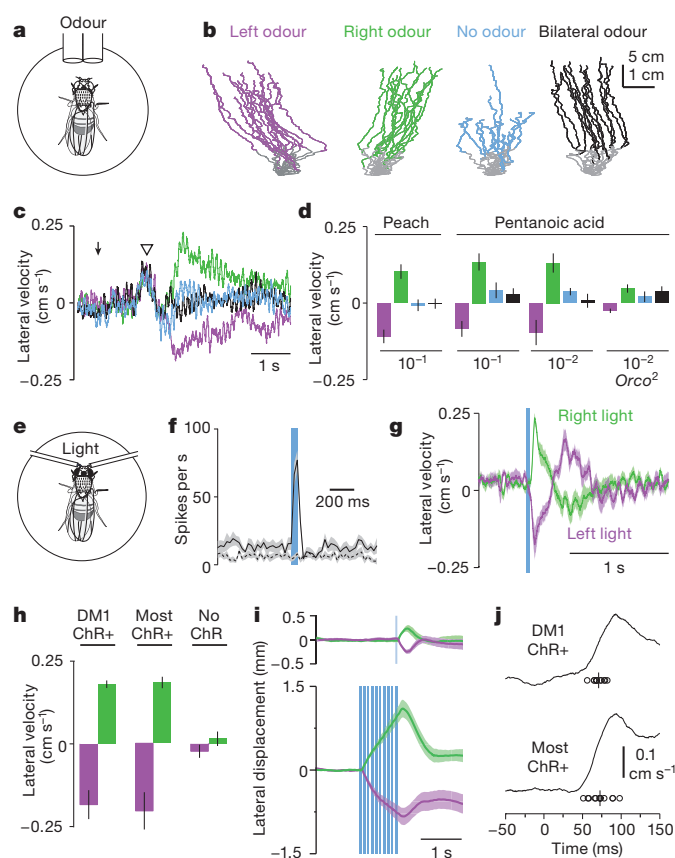and so this result implies that asymmetric input to strictly bilateral glomeruli can produce turning.

To target a single bilateral ORN type more selectively, we used an optogenetic approach (Fig. 1e). We expressed channelrhodopsin-2 (ChR2) in the ORNs that express the olfactory receptor Or42b and project to glomerulus DM1. These ORNs fired a burst of 3 or 4 spikes during a 50-ms light pulse directed at the antenna with a fine fibreoptic filament (Fig. 1f). In the same fly, light had no effect on adjacent ORNs that did not express ChR2 (Fig. 1f). Illuminating one antenna produced a turn towards the stimulus (Fig. 1g and Supplementary Fig. 2). Conversely, light offset produced a compensatory turn in the other direction (Fig. 1g, i and Supplementary Note 1). The turns evoked by ChR2 in DM1 ORNs were as large as the turns evoked by ChR2 in most ORN types (under the control of the *Orco-Gal4* line; Fig. 1h). Thus, bilateral ORNs can support lateralization behaviour, and turning can be elicited by just a few spikes per neuron in one ORN type.

The temporal control permitted by the optogenetic approach enabled us to precisely determine the latency of the behavioural response. Notably, turning began within approximately 70 ms of the ORN response onset (Fig. 1j). This is faster than the fly's stride period[14] (about 100 ms, Supplementary Fig. 1). Thus, there must be a rapid mechanism in the brain for extracting lateralized information from sensory neurons.

We next asked how asymmetric odour stimuli are encoded at the level of PN spikes. We removed one antenna in order to lateralize the odour stimulus, and we made simultaneous cell-attached recordings from PNs ipsilateral and contralateral to the intact antenna (Fig. 2a, b). We recorded from pairs of PNs in the same glomerulus on different sides of the brain ('sister PNs'), using green fluorescent protein (GFP) to target our electrodes to PNs in glomerulus DM6 or DM1. We used odours that preferentially activate either DM6 ORNs or DM1 ORNs (pentanoic acid or dilute ethyl acetate[12]). In these experiments, we found a small but consistent asymmetry in PN odour responses. This was apparent in the latency to the first odour-evoked spike: the first ipsilateral spike occurred 2.47 ± 0.70 ms earlier than the first contralateral spike in DM6 PNs, and 1.01 ± 0.41 ms earlier in DM1 PNs ($n = 16$ and 6; Fig. 2c). The latency difference between ipsilateral and contralateral spikes was statistically significant for DM6 PN pairs, but this difference was not significant for DM1 PN pairs ($P < 0.005$ and $P = 0.06$, respectively; Wilcoxon signed-rank tests). In addition, we found an asymmetry in odour-evoked firing rates: ipsilateral firing rates were on average about 50% higher than contralateral firing rates for DM6 PNs, with a larger asymmetry for DM1 PNs (Fig. 2d–g). The asymmetry was observed even during the spontaneous firing of these cells, and it was proportionately similar for all odour concentrations (Fig. 2d, g). The asymmetry in spikes must be due to an asymmetry in synaptic currents. Indeed, asymmetric stimulation of the antennae produced systematically larger inward currents in ipsilateral versus contralateral PNs (Supplementary Fig. 3).

Synaptic currents in PNs reflect the combined effects of ORNs and local neurons, most of which release GABA (γ-aminobutyric acid)[15–17].

[1]Department of Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, Massachusetts 02115, USA. [2]Rowland Institute, Harvard University, 100 Edwin Land Boulevard, Cambridge, Massachusetts 02142, USA. [3]Department of Organismic and Evolutionary Biology and Center for Brain Science, Harvard University, 52 Oxford Street, Cambridge, Massachusetts 02138, USA.

Figure 1 | *Drosophila* can lateralize odours based on bilateral receptor-neuron input to a single pair of glomeruli. a, Schematic diagram of a fly on a spherical treadmill, with odour tubes directed at each antenna.
b, Representative running trajectories from a single fly. Each trace is a different trial lasting 20 s. The grey portion of each trace indicates the pre-odour baseline period. Flies turn towards lateralized odour (fermented peach extract), but otherwise tend to run straight. **c**, Time course of mean lateral velocity in olfactory stimulation experiments ($n = 9$ flies). Positive values denote rightward turns and negative values denote leftward turns. Arrow indicates the onset of air flow through the tubing. The open arrowhead shows where clean air from the tip of the olfactometer first reaches the flies; this elicits a rightward turn that reflects either a systematic asymmetry in the tethering of the flies or an inherent handedness in the flies. Once this clean air is evacuated, odorized air elicits asymmetrical turning in the fly. Oscillations in lateral velocity are caused by the fly's stride rhythm (Supplementary Fig. 1). Colours as in **b**. **d**, Mean lateral velocity ($\pm$ s.e.m.) is significantly different for right versus left odour (green versus magenta bars). This was true for peach odour (10^{-1} dilution, $P < 0.005$, $n = 9$ flies, Wilcoxon signed-rank test). It was also true for pentanoic acid (10^{-1} and 10^{-2} dilutions, $n = 8$ and 12, $P < 0.01$ and $P < 0.005$, Wilcoxon signed-rank tests). In *Orco²* mutant flies, responses to right and left odour were still significantly different ($n = 12$ flies, $P < 0.005$, Wilcoxon signed-rank test) but were much smaller. The average difference between left and right (computed within each fly) is significantly smaller in *Orco²* flies compared to control flies ($P < 0.05$, Mann–Whitney *U*-test). The residual response in the mutants is probably due to the ORNs that do not rely on *Orco* (refs 11, 13). **e**, Schematic diagram of a fly on a spherical treadmill, with a fibre-optic light guide directed at each antenna. **f**, Time course of mean light-evoked firing rate in DM1 ORNs (ChR+, solid line) versus other ORN types (ChR−, dashed line). Shaded bands here and elsewhere represent $\pm$ s.e.m. **g**, Time course of mean lateral velocity during optogenetic stimulation in flies in which DM1 ORNs are ChR2+ **h**, Mean lateral velocity is significantly different for left and right antennal illumination in flies in which DM1 ORNs are ChR2+, or in flies in which most ORNs are ChR2+ ($n = 10$ and 12, $P < 0.005$ and $P < 5 \times 10^{-4}$, Wilcoxon signed-rank tests), but not in flies in which no ORNs express ChR2 ($P = 0.13$, $n = 10$). To express ChR2 in most ORNs, we used the *Orco-Gal4* line. Flies that lack ChR2 expression have the *UAS-ChR2* transgene but no Gal4 transgene. **i**, Time course of mean lateral displacement in response to a single light pulse (top) or a train of pulses (bottom). After the train, the magnitude of the compensatory turn is larger ($P < 0.005$, *t*-test), but the flies are also significantly less accurate in returning to their original running trajectory ($P < 0.005$, *t*-test). **j**, Time course of absolute mean lateral velocity on an expanded scale around the time of light onset (0 ms). Open circles show the turning latency for each fly. The mean latency across flies is shown as a vertical bar.

We asked whether the asymmetry in PN activity requires GABAergic inhibition, by bath-applying GABA$_A$ and GABA$_B$ receptor antagonists. These antagonists elevated PN firing rates, but the difference between ipsilateral and contralateral PN firing rates was not significantly altered (Fig. 3a, b).

Our results suggest that the asymmetry originates at the level of ORN input to PNs. To compare ORN–PN synapses in ipsilateral versus contralateral PNs, we made simultaneous whole-cell recordings of spontaneous excitatory postsynaptic currents (sEPSCs) in pairs of sister PNs. These sEPSCs are known to arise from ORN–PN synapses[9,10]. Each ORN–PN synapse consists of many release sites with high release probability, and spike-evoked synaptic events are therefore large and reliable[10]. In recordings from sister PNs in glomerulus DM6, almost all sEPSCs occurred in a paired fashion (Fig. 3c). Almost 100% of the sEPSCs in one cell had a corresponding paired event in the other cell, and this percentage was not significantly different in ipsilateral versus contralateral PNs (98.3 ± 0.6% versus 96.4 ± 1.3%, $n = 15$ pairs, $P = 0.11$; Wilcoxon signed-rank test). This result indicates that the number of unitary ORN–PN synaptic connections is essentially identical on the ipsilateral and contralateral sides, and it also indicates that action-potential conduction failures do not occur to any notable degree in the contralateral axon branches of ORNs, although we cannot completely exclude the idea that this occurs at a low rate.

These dual whole-cell recordings enable us to measure the mean difference in ipsilateral versus contralateral sEPSC arrival times (0.80 ± 0.51 ms). The difference in arrival times represents the delay imposed by axonal conduction between sister glomeruli. This delay is less than the delay between the first contralateral odour-evoked spike and the first ipsilateral odour-evoked spike in glomerulus DM6 (2.47 ms, see above). This makes sense, because the difference in first spike times is the result of two delays: the delay imposed by axonal conduction and the delay that results from less net synaptic excitation in the contralateral PN (Supplementary Fig. 3), which means that

the contralateral PN requires more integration time to reach spike threshold.

Comparing the amplitude of paired sEPSCs on an event-by-event basis, we found considerable variation in the relative amplitude of ipsilateral versus contralateral EPSCs (Fig. 3d). This is expected, given that synaptic vesicle release is a stochastic process occurring independently at ipsilateral and contralateral synapses. When many sEPSCs were considered together, there was a consistent asymmetry in the amplitude of sEPSCs (Fig. 3d). Averaged across all experiments, ipsilateral sEPSCs were 39% larger in amplitude than their contralateral counterparts (Fig. 3e). Thus, although ORN spikes reliably invade both ipsilateral and contralateral axon branches, each ORN spike typically has a stronger effect on ipsilateral PNs.

In principle, the asymmetry in sEPSC amplitudes could have either a pre- or postsynaptic locus. To determine whether we can observe this asymmetry at the level of presynaptic release sites, we expressed synaptobrevin–GFP (a marker of presynaptic release sites) in DM6 ORNs. We removed one antenna and allowed 3 days for the cut ORN axons to degenerate, leaving only the axons from the intact antenna. We found that total synaptobrevin fluorescence was on average 41 ± 16% higher on the ipsilateral side, and the ipsilateral:contralateral ratio of synaptobrevin fluorescence was significantly greater than 1 ($P < 0.05$, $n = 5$, *t*-test, Fig. 4a). This suggests an asymmetry in the number or size of neurotransmitter release sites. This is consistent with a previous study reporting that a plasma membrane marker
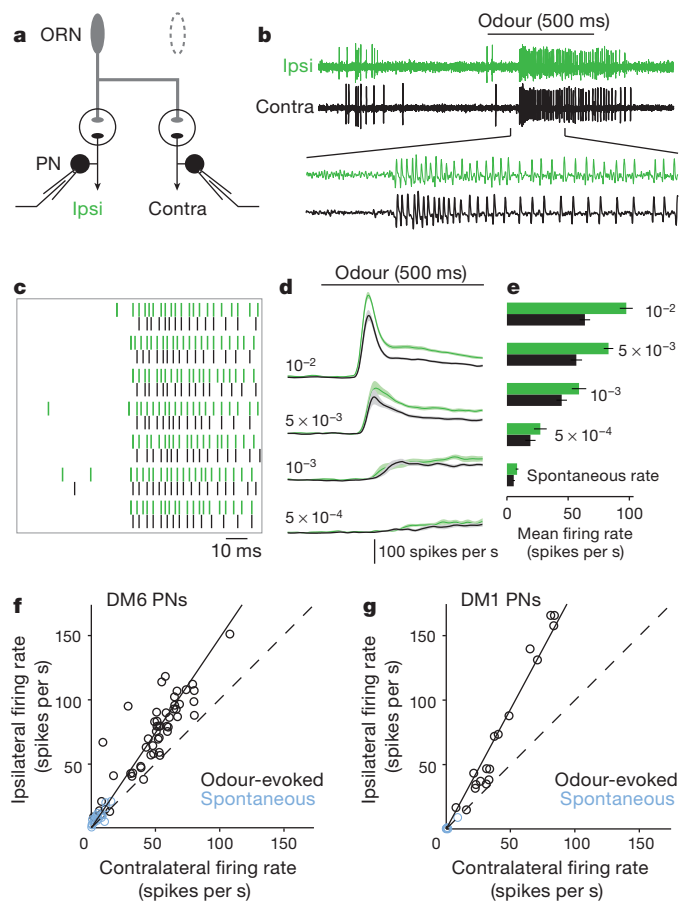
Figure 2 | Lateralized odours produce an asymmetry in spike latency and spike rate in antennal lobe PNs. a, Schematic of the set up for simultaneous cell-attached recording. An ORN axon innervates a glomerulus on each side of the midline, where it synapses onto postsynaptic PNs. In these experiments, one antenna was removed to lateralize the odour. **b**, Sample cell-attached recordings from a pair of PNs (postsynaptic to glomerulus DM6). The odour stimulus is a 500-ms pulse of pentanoic acid (10^{-2} dilution). Enlarged segment below is 250 ms. **c**, Raster plot showing the spiking responses of ipsilateral and contralateral DM6 PNs (odour is pentanoic acid at 10^{-2} dilution). Each pair of rows represents a single trial from the same pair of neurons. The raster starts 180 ms after the nominal odour-pulse onset. Note the shorter ipsilateral latency (see text). **d**, Time course of mean firing rates of DM6 PNs to a descending series of pentanoic acid concentrations. **e**, Mean firing rates of DM6 PNs ($\pm$ s.e.m.) in response to pentanoic acid concentrations. Spontaneous rates are also shown. Mean rate is computed over a 500-ms period starting 100 ms after the olfactometer is activated ($n = 16, 15, 10, 10$ and 21 pairs, in descending order of concentration). **f**, Trial-averaged ipsilateral versus contralateral firing rates for DM6 PNs. Several odour concentrations were used in each experiment, and each point represents a different experiment-concentration combination. Note that the ipsilateral–contralateral difference is present in spontaneous activity (blue symbols). Significance was assessed by fitting a line through the origin for each individual experiment; these slopes were significantly different from unity (mean slope = 1.47 ± 0.09, $P < 10^{-4}$, Mann–Whitney U-test, $n = 21$ experiments). The dashed line is unity; the solid line is the mean of all linear fits in the individual experiments. **g**, Same as **f**, but for DM1 PNs (mean slope = 1.86 ± 0.04, $P < 0.05$, Mann–Whitney U-test, $n = 6$ experiments). Odour stimuli were ethyl acetate at 10^{-12}, 10^{-10} and 10^{-6} dilutions. In DM1 PNs, spontaneous firing rates are close to zero.

was more abundant on the ipsilateral side[18]. In that study, 1 antenna was severed and imaging was performed 3 days later, as here. Functional remodelling can occur in that time period[18,19], so both findings should be interpreted cautiously.

Next, to visualize presynaptic calcium, we expressed the genetically encoded calcium indicator GCaMP3.0 in ORNs (using the *pebbled-Gal4*
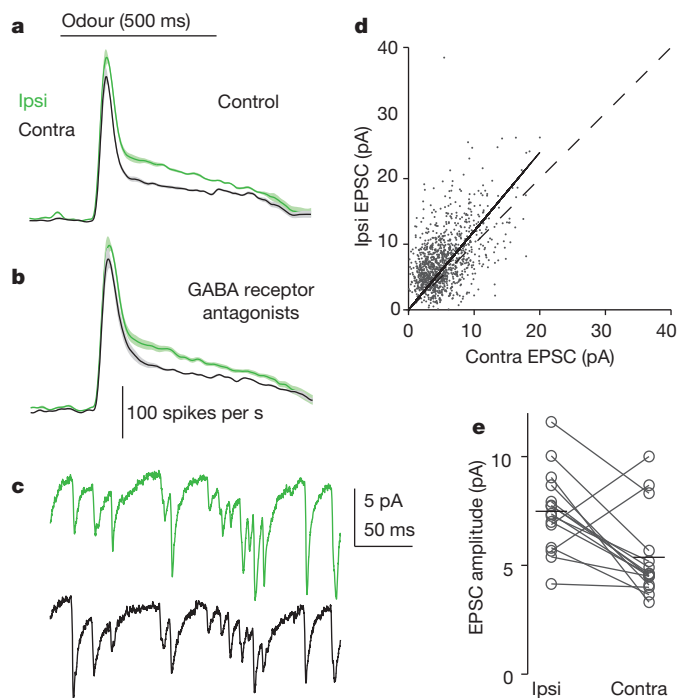
Figure 3 | The asymmetry arises at the level of ORN–PN synapses. a, Mean DM6 PN responses to pentanoic acid (10^{-2} dilution) in normal saline. **b**, Same as **a**, but with the addition of the $GABA_A$ receptor antagonist picrotoxin (5 µM) and the $GABA_B$ receptor antagonist CGP54626 (50 µM). Ipsilateral firing rates are significantly higher than contralateral rates ($P < 0.01$), and the antagonists have no significant effect on the ipsilateral–contralateral difference ($P = 0.86$, $n = 10$, two-way analysis of variance (ANOVA) on data from **a** and **b**). The ipsilateral–contralateral difference in peak firing rates becomes somewhat smaller, but because this occurs only at the peak, it is probably due to the near-saturation of PN firing rates. **c**, Whole-cell recordings of sEPSCs from a pair of DM6 PNs. The PN ipsilateral to the intact antenna is in green. **d**, Ipsilateral versus contralateral EPSC amplitudes in a typical pair of DM6 PNs. Each point represents a pair of sEPSCs ($n = 1,213$). Dashed line is unity; solid line is a linear fit constrained to intersect the origin. **e**, Group data showing mean sEPSC amplitudes in all pairs of DM6 PNs. Each symbol is a different experiment and horizontal lines represent means across experiments. sEPSC amplitudes are significantly larger in ipsilateral PNs ($n = 15$ pairs, $P < 0.05$, Wilcoxon signed-rank test).

line) and used two-photon microscopy to visualize calcium signals in ORN axon terminals. To lateralize the odour, we removed one antenna immediately before the experiment (Fig. 4b). We used pentanoic acid to evoke a fluorescence increase preferentially in glomerulus DM6 (Fig. 4c). We found that the size of the calcium response was significantly greater on the side of the brain ipsilateral to the intact antenna (Fig. 4c, d). This asymmetry did not require feedback from central circuits, because it persisted after washing in mecamylamine (to block nicotinic acetylcholine receptors, which mediate ORN–PN synaptic transmission[10]) along with picrotoxin and CGP54626 hydrochloride ($GABA_A$ and $GABA_B$ receptor antagonists, respectively) (Fig. 4e). In separate experiments, we saw similarly asymmetric calcium signals in glomerulus DL5, using an odour stimulus that is relatively selective for DL5 ORNs[20] (trans-2-hexenal, 10^{-5} dilution, data not shown). We also saw a similar asymmetry in presynaptic currents, using simultaneous bilateral field potential recordings (Supplementary Fig. 4). Together, these results demonstrate that the asymmetry in EPSC amplitudes has a presynaptic origin.

Interestingly, we found a roughly equal asymmetry (~40%) in the amplitude of sEPSCs, the level of synaptobrevin fluorescence, and the odour-evoked GCaMP3 fluorescence change. The simplest explanation for these results is that the ipsilateral arbor is 40% larger than the contralateral arbor. If everything else is equal, then this mechanism
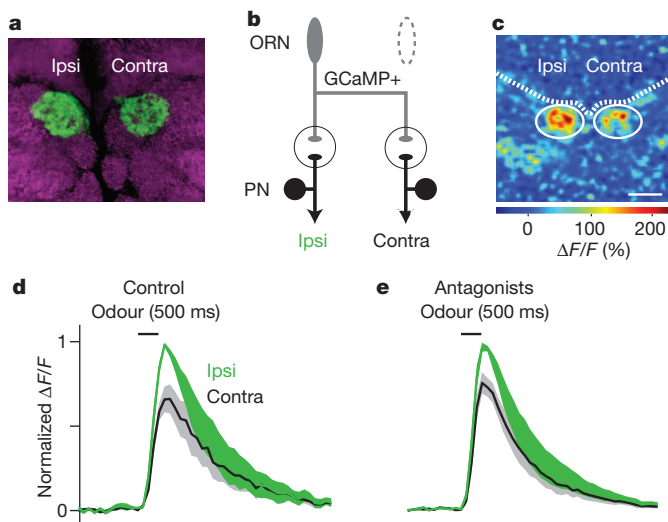
Figure 4 | The asymmetry in ORN–PN synapses has a presynaptic origin.
a, Synaptobrevin–GFP (green) was expressed in DM6 ORNs to label neurotransmitter release sites, and 1 antenna was removed 3 days before to allow the cut axons to degenerate. The brain is viewed from the anterior face. Magenta shows neuropil contours (nc82 immunofluorescence). Scale bar, 10 μm. **b**, Schematic of calcium-recording configuration. ORNs express GCaMP, and one antenna was removed to lateralize the odour. **c**, Changes in fluorescence in ORN axon terminals in response to pentanoic acid (10^{-2} dilution). Solid circles outline the DM6 glomeruli, and the dashed line shows the anterior boundary of the antennal lobe neuropil. The brain is viewed from the dorsal side. Scale bar, 10 μm. **d**, Mean time course of odour-evoked calcium signals in ORN axons innervating glomerulus DM6. Black bar indicates the timing of the odour stimulus (pentanoic acid at 10^{-2} dilution, 500 ms). **e**, Same as **d**, except after adding picrotoxin, CGP54626, and the nicotinic receptor antagonist mecamylamine (200 μM). Ipsilateral responses are significantly larger than contralateral responses ($P < 0.05$) and the antagonists have no significant effect on the ipsilateral–contralateral difference ($P = 0.75$, $n = 6$, two-way ANOVA on data from panels **d** and **e**).

should produce a proportional change in all these measurements (Supplementary Note 2).

Our results are consistent with the finding that in flies with one antenna removed, 2-deoxyglucose uptake is higher on the side ipsilateral to the intact antenna[21]. A previous study using calcium imaging has reported that asymmetric ORN stimulation can produce stronger signals in ipsilateral PNs than in contralateral PNs[22], which is also broadly consistent with our results. However, in contrast to our results, this asymmetry was restricted to just a few glomeruli. This finding is puzzling, given that a variety of stimuli can produce turning behaviour (Fig. 1 and refs 3–5). The same study also concluded that GABAergic inhibition mediates the ipsilateral–contralateral asymmetry in PNs[22]. These discrepancies may reflect limitations of the imaging techniques used in that study, as well as differences between stimuli or glomeruli.

Another previous study failed to find any significant ipsilateral–contralateral differences in the strength of ORN–PN synapses[10]. However, the measurements in the study involved sequential recordings from sister PNs, rather than simultaneous recordings, and the measurements were of EPSCs arising from a single ORN per PN. Here we have better statistical power because we have made simultaneous recordings from sister PNs, and because we sampled sEPSCs arising from many ORNs. Our finding that ipsilateral and contralateral synapses differ only modestly explains why this difference has been difficult to resolve previously.

It is well-known that a single axon can form neurotransmitter release sites with different properties on different postsynaptic cell types (for example, principal neurons versus interneurons[23]). Here we have shown that ORN axons discriminate between two classes of target cells (left and right) that are morphologically and functionally

identical, and that share the same lineage and birth dates[24]. This is reminiscent of the circuitry of the leech midbody ganglion, in which individual mechanoreceptor axons make stronger synapses onto ipsilateral versus contralateral sister cells that are otherwise functionally identical[25]. It is possible that these target cells may be distinguished because right and left sister neurons express different molecular tags[26], thereby allowing a given axon to recognize them as ipsilateral versus contralateral in relation to itself. Alternatively, an axon might form more release sites at proximal locations than at distal locations; this cell-intrinsic mechanism would suffice because contralateral sites are always more distal than ipsilateral sites.

Our results reveal that even small signals in the *Drosophila* nervous system can be behaviourally relevant. The stimuli we used in our optogenetic experiments produced only a slight fractional difference in input to the right and left sides of the brain. This difference amounted to 5% over the 50-ms stimulus period (Supplementary Note 3). The finding that this incremental difference is relevant for behaviour should provide an additional motive for the continued development of sensitive methods for monitoring neural activity in the fly brain.

METHODS SUMMARY

The odour-delivery device used for the olfactory behavioural experiments was specially designed to deliver no lateralized mechanical cues, and we carried out control experiments to confirm that no turning was observed when no odour was present (Supplementary Fig. 5). The spherical treadmill apparatus was constructed by floating a small plastic sphere on a jet of compressed air. We measured the forward velocity (pitch) and lateral velocity (roll) of the sphere in the apparatus by placing the sensor from an optical mouse underneath the sphere. Photostimulation of *Drosophila* ORNs was achieved by butt-coupling a blue light-emitting diode (LED) to a fibre-optic filament (50 μm diameter) and positioning the tip of the fibre approximately 150 μm away from the fly's antenna. In all behavioural experiments using light to stimulate ORNs, the eyes and ocelli of the fly were shielded from light by painting them with ink. *In vivo* extracellular recordings from ORNs and patch-clamp recordings from PNs were carried out as described previously[9,27]. In all electrophysiology or calcium-imaging experiments (except ORN recordings), one antenna was removed just before the experiment by a person who was not the experimenter, and the experimenter remained blind to which side of the brain was ipsilateral to the intact antenna. All analysis was also carried out blind to which side was ipsilateral. Calcium-imaging experiments were performed on a custom-built two-photon microscope. All aggregated data represent mean ± s.e.m. computed across experiments. See Supplementary Methods for details on the spherical treadmill, odour delivery, optogenetic stimuli and analysis of sEPSCs.

Full Methods and any associated references are available in the online version of the paper.

1. Stocker, R. F., Lienhard, M. C., Borst, A. & Fischbach, K. F. Neuronal architecture of the antennal lobe in *Drosophila melanogaster*. *Cell Tissue Res.* **262**, 9–34 (1990).
2. Couto, A., Alenius, M. & Dickson, B. J. Molecular, anatomical, and functional organization of the *Drosophila* olfactory system. *Curr. Biol.* **15**, 1535–1547 (2005).
3. Borst, A. & Heisenberg, M. Osmotropotaxis in *Drosophila melanogaster*. *J. Comp. Physiol. A* **147**, 479–484 (1982).
4. Duistermars, B. J., Chow, D. M. & Frye, M. A. Flies require bilateral sensory input to track odor gradients in flight. *Curr. Biol.* **19**, 1301–1307 (2009).
5. Flugge, C. Geruchliche raumorientierung von *Drosophila melanogaster*. *Z. Vgl. Physiol.* **20**, 463–500 (1934).
6. Kennedy, J. S. & Moorhouse, J. E. Laboratory observations on locust responses to wind-borne grass odour. *Entomol. Exp. Appl.* **12**, 487–503 (1969).
7. Martin, H. Osmotropotaxis in the honey-bee. *Nature* **208**, 59–63 (1965).
8. Hangartner, W. Spezifität und inaktivierung des spurpheromons von *Lasius fuliginosus* (Latr.) und orientierung der arbeiterinnen im duftfeld. *Z. Vgl. Physiol.* **57**, 103–136 (1967).
9. Kazama, H. & Wilson, R. I. Origins of correlated activity in an olfactory circuit. *Nature Neurosci.* **12**, 1136–1144 (2009).
10. Kazama, H. & Wilson, R. I. Homeostatic matching and nonlinear amplification at genetically-identified central synapses. *Neuron* **58**, 401–413 (2008).
11. Silbering, A. F. *et al.* Complementary function and integrated wiring of the evolutionarily distinct *Drosophila* olfactory subsystems. *J. Neurosci.* **31**, 13357–13375 (2011).

12. Hallem, E. A. & Carlson, J. R. Coding of odors by a receptor repertoire. *Cell* **125,** 143–160 (2006).
13. Benton, R., Vannice, K. S., Gomez-Diaz, C. & Vosshall, L. B. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* **136,** 149–162 (2009).
14. Strauss, R. & Heisenberg, M. Coordination of legs during straight walking and turning in *Drosophila melanogaster*. *J. Comp. Physiol. A* **167,** 403–412 (1990).
15. Chou, Y. H. *et al.* Diversity and wiring variability of olfactory local interneurons in the *Drosophila* antennal lobe. *Nature Neurosci.* **13,** 439–449 (2010).
16. Okada, R., Awasaki, T. & Ito, K. Gamma-aminobuyric acid (GABA)-mediated neural connections in the *Drosophila* antennal lobe. *J. Comp. Neurol.* **514,** 74–91 (2009).
17. Das, A. *et al. Drosophila* olfactory local interneurons and projection neurons derive from a common neuroblast lineage specified by the empty spiracles gene. *Neural Dev.* **3,** 33 (2008).
18. Berdnik, D., Chihara, T., Couto, A. & Luo, L. Wiring stability of the adult *Drosophila* olfactory circuit after lesion. *J. Neurosci.* **26,** 3367–3376 (2006).
19. Kazama, H., Yaksi, E. & Wilson, R. I. Cell death triggers olfactory circuit plasticity via glial signaling in *Drosophila*. *J. Neurosci.* **31,** 7619–7630 (2011).
20. Olsen, S. R., Bhandawat, V. & Wilson, R. I. Divisive normalization in olfactory population codes. *Neuron* **66,** 287–299 (2010).
21. Rodrigues, V. Spatial coding of olfactory information in the antennal lobe of *Drosophila melanogaster*. *Brain Res.* **453,** 299–307 (1988).
22. Agarwal, G. & Isacoff, E. Specializations of a pheromonal glomerulus in the *Drosophila* olfactory system. *J. Neurophysiol.* **105,** 1711–1721 (2011).
23. Pelkey, K. A. & McBain, C. J. Differential regulation at functionally divergent release sites along a common axon. *Curr. Opin. Neurobiol.* **17,** 366–373 (2007).
24. Jefferis, G. S., Marin, E. C., Stocker, R. F. & Luo, L. Target neuron prespecification in the olfactory map of *Drosophila*. *Nature* **414,** 204–208 (2001).
25. Lockery, S. R. & Kristan, W. B. Jr. Distributed processing of sensory information in the leech. II. Identification of interneurons contributing to the local bending reflex. *J. Neurosci.* **10,** 1816–1829 (1990).
26. Chintapalli, V. R. *et al.* Functional correlates of positional and gender-specific renal asymmetry in *Drosophila*. *PLoS ONE* **7,** e32577 (2012).
27. Bhandawat, V., Olsen, S. R., Schlief, M. L., Gouwens, N. W. & Wilson, R. I. Sensory processing in the *Drosophila* antennal lobe increases the reliability and separability of ensemble odor representations. *Nature Neurosci.* **10,** 1474–1482 (2007).

Author Contributions Q.G. and R.I.W. designed the experiments. Q.G. carried out all of the experiments, except for the calcium imaging, which was performed by Q.G. and E.J.H., and the synaptobrevin imaging, which was performed by R.I.W. Q.G. analysed the data. J.K. and B.L.d.B. helped to design and build the spherical treadmill apparatus. Q.G. and R.I.W. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.I.W. (rachel_wilson@hms.harvard.edu).

METHODS

For all electrophysiology and calcium-imaging experiments, each stimulus was presented in 4 to 6 trials, and the responses were averaged together to produce the response for that experiment. Before statistical tests, all the data in a given panel were collectively tested for normality using a Shapiro–Wilk test. If the data were not normally distributed, then a non-parametric test was used for all the comparisons in that panel.

Fly stocks. Flies (1 to 2 days post eclosion) were raised on conventional cornmeal medium at 25 °C with a 12 h light–dark cycle. For behavioural experiments, we used a fly stock recently derived from wild-caught *Drosophila melanogaster* (provided by M. Dickinson), because these flies showed more robust running than inbred stocks. For optogenetic experiments, we used flies with one copy of the *Gal4* transgene and two copies of the *UAS-ChR2–YFP* transgene (one on chromosome 2 and one on chromosome 3). All flies harbouring the *UAS-ChR2–YFP* transgene, including control flies lacking the *Gal4* transgene, were raised on food containing all-trans retinal. All-trans retinal was prepared as a stock solution in ethanol (35 mM), and 10 μl of this stock was mixed into approximately 5 ml of rehydrated potato flakes and added to the top of a vial of conventional food. DM6 PN recordings were carried out using the genotype *NP3062-Gal4,UAS-CD8–GFP*. DM1 PN recordings were carried out using the genotype *NP5221-Gal4;UAS-CD8–GFP*. Synaptobrevin-imaging experiments were carried out in the genotype *Or67d-Gal4/UAS-n-syb–eGFP*. Calcium-imaging experiments were carried out in the genotype *pebbled-Gal4;UAS-GCAMP3.0*. Fly stocks were previously published as follows: wild-caught flies[28], *Or42b-Gal4* (ref. 29, Bloomington no. 9971), *Orco-Gal4* (*Or83b-Gal4*, ref. 30, Bloomington no. 26818), *UAS-ChR2–YFP* (ref. 31), *NP3062-Gal4* (ref. 10, DGRC no. 113083), *NP5221-Gal4* (ref. 32, DGRC no. 104906), *UAS-CD8–GFP* (ref. 33, Bloomington no. 5136 and no. 7465), *Or67d-Gal4* (ref. 2, Bloomington no. 23904), *UAS-n-syb–eGFP* (ref. 34, Bloomington no. 6921), *pebbled-Gal4* (ref. 35), *UAS-GCaMP3.0* (ref. 36, Bloomington no. 32116), *Orco*[2] (ref. 30, Bloomington no. 23130). For details on the spherical treadmill, odour delivery and optogenetic stimulation, see Supplementary Methods.

ORN recordings. Extracellular recordings from ORNs were conducted as described previously[2]. In brief, the fly was mounted on the end of a cut pipette tip and secured using paraffin wax. The third antennal segment was restrained using hooks fabricated from glass capillaries positioned with micromanipulators. Specific ORN types were identified based on the size and location of the sensillum, together with the action potential wave forms and odour responses of the ORNs. ChR2-expressing ORNs were stimulated using the same fibre-optic apparatus used in the behavioural experiments. To verify that the light from the fibre-optic cables was sufficiently focused, we also recorded from ChR2-expressing ORNs in the antenna contralateral to the illumination, and we confirmed that these neurons showed no responses to light (data not shown). The bins in Fig. 1f are 25-ms wide and are plotted versus the time that corresponds to the end of each bin.

PN recordings. Flies were dissected as described previously[27], except that the odour source was lateralized by removing one antenna just before recording. The identity of the intact antenna was pseudo-randomized between preparations. We targeted our electrodes to particular PNs by expressing GFP in these PNs (see the section on fly stocks above). The identity of each recorded PN was confirmed using a panel of diagnostic odours, and was corroborated in pilot studies using immunohistochemistry to identify the glomerulus innervated by the PN dendrites. Data were low-pass filtered at 1 kHz for cell-attached recordings and at 5 kHz for whole-cell recordings using an Axopatch 200B amplifier (Molecular Devices) and digitized at 10 kHz. The external saline was prepared as described previously[27] and was continuously bubbled with a blend of 95% O_2 and 5% CO_2. Whole-cell experiments were carried out in voltage-clamp mode with a holding potential of −60 mV. The internal solution contained (in mM) caesium aspartate, 140; HEPES, 10; MgATP, 4; Na_3GTP, 0.5; EGTA, 1; CsCl, 1; biocytin hydrazide, 13

(pH 7.3, osmolarity adjusted to ~265 mOsm). Cell-attached recordings were carried out in voltage-clamp mode and the command potential was adjusted so that the amplifier did not pass any current. For cell-attached recordings, we used patch pipettes filled with either external saline or the internal pipette solution used for the whole-cell recordings. As both DM6 PNs and DM1 PNs have low spontaneous firing rates, the identification of the first spike in the odour response was generally unambiguous. To ensure that the first-spike latency was measured in an unbiased way, we used an automated algorithm to identify the first spike after odour onset; this was defined as the spike preceding the first inter-spike interval that is shorter than 90% of the inter-spike intervals in the pre-odour period. For details on the analysis of sEPSCs, and on synaptobrevin-imaging, see Supplementary Methods.

Calcium imaging. We expressed the genetically encoded calcium indicator GCaMP3.0 in olfactory receptor neurons using the Gal4/UAS system. We used flies homozygous for both *UAS-GCaMP3.0* and *pebbled-Gal4*, which drives expression in most or all ORNs (along with other antennal neurons that do not project to the antennal lobe). These flies were older than those used in the rest of the study (25 to 35 days post eclosion) because we found that this increased GCaMP3.0 fluorescence. Just before each experiment, one antenna was removed to lateralize the odour. The olfactometer was identical to the one used for electrophysiology experiments. The inter-trial interval was 45 s. Data were collected on a custom-built two-photon microscope at a frame rate of 7.8 Hz. Images were collected at each of several *z* planes that collectively spanned the depth of the DM6 glomeruli, and all these images were averaged together and filtered with a two-dimensional Gaussian (s.d. = 5 pixels). The odour-evoked increase in fluorescence was determined by drawing a region of interest around the activated glomeruli in the sequence of raw fluorescence images. We then calculated the change in fluorescence divided by the baseline fluorescence ($\Delta F/F$) within this region for each frame. The peak calcium signal was measured from odour onset through the next 5 frames (640 ms total). All imaging experiments and analyses were carried out blind with regard to which antenna was removed. Resting fluorescence (F) was not significantly different in the ipsilateral and contralateral glomeruli. Although there are probably more release sites on the ipsilateral side (Fig. 4a), any asymmetry in resting fluorescence resulting from spontaneous activity is evidently too small to resolve. In these experiments, we did not detect a fluorescence increase outside of the expected glomeruli, except for a relatively weak and intermittent signal in DM1; however, DM1 also responded when replaced the odour vial with an empty vial, indicating that it was responding to a contaminant in our delivery system rather than the odour.

28. Bhandawat, V., Maimon, G., Dickinson, M. H. & Wilson, R. I. Olfactory modulation of flight in *Drosophila* is sensitive, selective and rapid. *J. Exp. Biol.* **213,** 3625–3635 (2010).
29. Fishilevich, E. & Vosshall, L. B. Genetic and functional subdivision of the *Drosophila* antennal lobe. *Curr. Biol.* **15,** 1548–1553 (2005).
30. Larsson, M. C. *et al.* Or83b encodes a broadly expressed odorant receptor essential for *Drosophila* olfaction. *Neuron* **43,** 703–714 (2004).
31. Hwang, R. Y. *et al.* Nociceptive neurons protect *Drosophila* larvae from parasitoid wasps. *Curr. Biol.* **17,** 2105–2116 (2007).
32. Tanaka, N. K., Awasaki, T., Shimada, T. & Ito, K. Integration of chemosensory pathways in the *Drosophila* second-order olfactory centers. *Curr. Biol.* **14,** 449–457 (2004).
33. Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* **22,** 451–461 (1999).
34. Zhang, Y. Q., Rodesch, C. K. & Broadie, K. Living synaptic vesicle marker: synaptotagmin-GFP. *Genesis* **34,** 142–145 (2002).
35. Sweeney, L. B. *et al.* Temporal target restriction of olfactory receptor neurons by Semaphorin-1a/PlexinA-mediated axon-axon interactions. *Neuron* **53,** 185–200 (2007).
36. Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods* **6,** 875–881 (2009).

ARTICLE

Towards an exact description of electronic wavefunctions in real solids

George H. Booth[1], Andreas Grüneis[1,2], Georg Kresse[2] & Ali Alavi[1]

The properties of all materials arise largely from the quantum mechanics of their constituent electrons under the influence of the electric field of the nuclei. The solution of the underlying many-electron Schrödinger equation is a 'non-polynomial hard' problem, owing to the complex interplay of kinetic energy, electron–electron repulsion and the Pauli exclusion principle. The dominant computational method for describing such systems has been density functional theory. Quantum-chemical methods—based on an explicit ansatz for the many-electron wavefunctions and, hence, potentially more accurate—have not been fully explored in the solid state owing to their computational complexity, which ranges from strongly exponential to high-order polynomial in system size. Here we report the application of an exact technique, full configuration interaction quantum Monte Carlo to a variety of real solids, providing reference many-electron energies that are used to rigorously benchmark the standard hierarchy of quantum-chemical techniques, up to the 'gold standard' coupled-cluster ansatz, including single, double and perturbative triple particle–hole excitation operators. We show the errors in cohesive energies predicted by this method to be small, indicating the potential of this computationally polynomial scaling technique to tackle current solid-state problems.

Although density functional theory has been the workhorse of computational materials science for several decades[1], systematic routes to improve the crucial but approximate exchange-correlation functionals do not exist[2]. In contrast, for molecular systems, a systematic hierarchy of approximate yet highly successful quantum-chemical techniques, such as coupled-cluster theory, has long been established[3]. This hierarchy has not yet been explored in solids, although initial implementations of its lower levels have been encouraging[4–11]. Part of the reason for this is the high computational complexity of quantum-chemical methods. The computational cost grows rapidly with the number of considered electrons N and with the basis set size M. Traditional full configuration interaction (FCI), that is, exact diagonalization, has combinatorial scaling and can cope with at most some ten electrons in a small basis; and even coupled-cluster methods, although requiring only a computational time that is polynomial in M and N, are extremely expensive. However, recent developments in methodology, as well as the increase in computer power, mean that it is now possible to address their accuracy and applicability in this domain.

In the development of quantum chemistry, FCI has played an invaluable benchmarking role, by providing exact results within a given basis. This has enabled the electron correlation problem to be addressed in isolation from other complicating factors inherent when comparing to experiment[12]. Moreover, FCI enables us to assess the degree to which the electronic wavefunction is dominated by a single determinant, and therefore which systems are likely to be well described by approximations such as many-body perturbation theory and coupled-cluster theory[13]. In solids, however, the absence of FCI means that the accuracy of such approximations cannot be easily gauged, especially in systems where correlations are expected to be strong. Here we provide, for the first time to our knowledge, FCI-quality energies in a range of realistic solids and unambiguously evaluate the accuracy of high-level quantum-chemical methods.

A further motivation for the application of quantum-chemical methods to solids comes from the multitude of recent developments, which hold the promise to reduce the computational cost beyond our present consideration. These include optimized virtual spaces[14], explicit correlation[15,16], exploitation of locality of correlation[5], and others[17–19], and should be directly transferable to the solid state. Combining high accuracy with increased efficiency, quantum-chemistry methods hold the promise to routinely bring high accuracy to computational materials science.

FCIQMC and the quantum-chemical hierarchy

The FCI quantum Monte Carlo (FCIQMC) method has emerged as a tool to calculate energies that are essentially identical to the true correlation energy captured by the basis set, whilst having a significantly lower computational scaling than a traditional brute force diagonalization of the problem (FCI). This makes it ideally suited for a systematic benchmarking of wavefunction-based methods in the solid state[20–25]. This method involves a stochastic sampling of a Slater determinant space constructed from the basis set—a function space of orthonormal antisymmetrized determinants in which the wavefunctions are expressed. This method has previously been applied to molecular systems[20–24] and the homogeneous electron gas[25], where energies were calculated that compared favourably, or in some cases surpassed in accuracy, those achieved with state-of-the-art diffusion Monte Carlo (DMC) techniques[26]. This provides the confidence to tackle more realistic solid-state systems here.

Although it is possible for DMC to be used as a benchmark for quantum-chemistry methods and vice versa[27], DMC does not operate in a Slater determinant space, but rather a real space representation of the wavefunction. As such, it would require quantum-chemical calculations to be converged to high accuracy with respect to the basis set size before any meaningful comparisons could be drawn. Instead, by comparing to values obtained in the same Slater determinant basis, robust comparisons between FCIQMC and more approximate diagrammatic methods can be drawn without the need for absolute convergence. More approximate methods can then also be used to

[1]University of Cambridge, Chemistry Department, Lensfield Road, Cambridge CB2 1EW, UK. [2]University of Vienna, Faculty of Physics and Center for Computational Materials Science, Sensengasse 8/12, A-1090 Vienna, Austria.

extrapolate to the infinite basis set limit[14,17,28]. Additionally, DMC requires an approximation for the nodal surface of the wavefunction. Although this error can be made relatively small[29,30], releasing the nodal surface is notoriously difficult for solids and greatly increases the computational demand. Similarly, another variant of QMC methods, auxiliary-field quantum Monte Carlo (AFQMC), although now operating in a space of Slater determinants and with favourable scaling, requires analogous constraints within the phaseless approximation in order to go to realistic system sizes and avoid transient energy estimates[31,32].

Here we illustrate the extension of the FCIQMC method to the solid state. The introduction of translational and crystal momentum symmetries, which arise from working with finite simulation cells with periodic boundary conditions, necessitates a change to the 'walker' dynamics and 'initiator' rules (see below for descriptions). In order to take advantage of these properties, the method is reformulated for complex wavefunctions. The presence of a 'phase' problem, rather than a simpler sign problem, is considered. A range of systems, from the extensively studied lithium hydride, to other ionic, covalent and rare gas solids is considered, while benchmarking the established quantum-chemical methods of second-order Møller–Plesset theory (MP2)[33], coupled-cluster singles and doubles (CCSD)[34], and the first implementation of perturbative triples (CCSD(T))[35] for periodic systems. These three methods are generally considered to possess a favourable trade-off between accuracy and cost, and are the most widely used of all quantum-chemical methods. Finally, we study the far more complex electronic structure of the charge-transfer solid NiO, to evaluate the ability of these quantum-chemical methods to handle strong correlation effects, where deficiencies are likely to be exposed.

Sampling in the solid state

The recently developed FCIQMC method[20–24] involves a discrete sampling of the wavefunction by signed 'walkers' which stochastically evolve within a Hilbert space of N-electron Slater determinants as illustrated in Fig. 1. The method converges rapidly with the number of walkers to the FCI limit, and generally the number of walkers required is a tiny fraction of the Hilbert space. This leads to a huge compression in the wavefunction information, while correctly reproducing exact time-averaged properties. The reason that a sufficient number of walkers is required has to do with overcoming the 'fermion sign problem' present in the Monte Carlo sampling of any fermionic
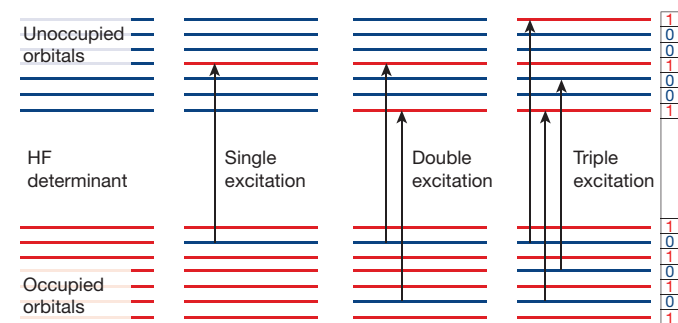

Figure 1 | FCIQMC in a nutshell. The antisymmetric many-electron wavefunction $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ is represented in a space of determinants constructed out of Hartree–Fock (HF) orbitals (occupied as well as unoccupied one electron wavefunctions, $\phi(\mathbf{r})$). Walkers are encoded by a bit string as shown on the right, where each bit corresponds to one orbital. '1' (red) and '0' (black) imply that the orbital is occupied and unoccupied, respectively. Because some Slater determinants (for instance, the Hartree–Fock determinant) have a very high occupation probability, additional bits are reserved to count the signed number of real and imaginary walkers on a determinant. The computational procedure involves selecting a new determinant from an existing walker with a certain sign and transition probability as discussed in the main text. The excitation number of a determinant refers to the numbers of holes which need to be introduced in the Hartree–Fock determinant to generate the determinant.

wavefunction[36]. The sign problem in this space manifests itself in the presence of a lower-energy state characterized by a combination of the $\pm\Psi$ degenerate solutions, which if not suppressed, leads to an exponential increase in noise. The growth of this state is controlled in FCIQMC by annihilation between walkers of opposite signs, which stabilizes the wavefunction to one signed solution[20,37]. The discrete space of Slater determinants allows FCIQMC to implement this annihilation exactly, and provided the walker population is large enough, will directly overcome the sign problem without requiring constraints on the wavefunction.

Because walker annihilation can only occur on occupied determinants, it is important to ensure that the newly occupied space remains sign-coherent to the currently sampled wavefunction. This is the rationale behind the 'initiator' rules used in the i-FCIQMC method[22], which is used exclusively in this work, whereby newly occupied determinants must have originated from a determinant with a population greater than a parameter n_{add}. By restricting the growth of the occupied space in this way, the walker density and hence annihilation rate is kept high, ensuring that propagation of noise in the system is kept to a minimum. This biases the dynamic in a small way, but rigorously converges onto exact energies of the Hamiltonian as the walker number increases.

The determinants in this work are composed from antisymmetrized products of one-electron orbitals obtained from a prior Hartree–Fock calculation in a large basis of periodic plane waves within the framework of the projector-augmented wave method, as implemented in VASP. If these orbitals are strictly real then the wavefunction Ψ can also be real[20]. In this work, however, the orbitals are complex Bloch functions, to account for the translational invariance of the potential. With these we can construct many body wavefunctions and use k-point sampling to ensure convergence, rather than sampling ever larger unit cells to remove finite-size effects.

Because it is necessary to correlate between sampled k-points, the number of explicitly correlated electrons and orbitals increases linearly with the number of sampled k-points yielding combinatorial scaling in the size of the Hilbert space (essentially exponential with the number of k-points). However, performing this sampling increases the number of zero Hamiltonian matrix elements between determinants, as crystal momentum must be conserved. By implementing an algorithm to stochastically generate only these momentum-allowed excitations, a saving that grows quadratically with the number of k-points is achieved, because both the accessible space is reduced, and the magnitude of the time step is increased.

To take advantage of these savings, it is necessary to work with complex orbitals, requiring a complementary set of both 'real' and 'imaginary' walkers in the FCIQMC dynamic, and a reformulation of the algorithm. The master equations of the FCIQMC method follow naturally from the imaginary-time Schrödinger equation, and are given by

$$-\frac{dN_\mathbf{i}}{d\tau} = (H_{\mathbf{ii}} - S)N_\mathbf{i} + \sum_{\mathbf{j} \neq \mathbf{i}} H_{\mathbf{ij}}N_\mathbf{j} \qquad (1)$$

where $N_\mathbf{i}$ represents the now complex walker population on determinant $D_\mathbf{i}$, τ represents imaginary time, S is a strictly real energy-offset parameter denoted the 'shift', which controls population growth, and $H_{\mathbf{ij}} = \langle D_\mathbf{i}|\hat{H}|D_\mathbf{j}\rangle$ is the many-electron Hamiltonian evaluated between two determinants. In each iteration, for each walker (real and imaginary) on a determinant $D_\mathbf{i}$, a suitable momentum-allowed excitation, $D_\mathbf{j}$, is generated. The real ($\Re$) and imaginary ($\Im$) parts of $H_{\mathbf{ij}}$ are considered in turn, and two attempts at generating new walkers on $D_\mathbf{j}$ are stochastically realized. For real parent walkers:

$$p_s^\Re(\mathbf{j}|\mathbf{i}) = \frac{\delta\tau|\Re[H_{\mathbf{ij}}]|}{p_{gen}(\mathbf{j}|\mathbf{i})}; \quad sign = -sign\left(\Re[N_\mathbf{i}]\Re[H_{\mathbf{ij}}]\right) \qquad (2)$$

$$p_s^{\Im}(\mathbf{j}|\mathbf{i}) = \frac{\delta\tau|\Im[H_{ij}]|}{p_{\mathrm{gen}}(\mathbf{j}|\mathbf{i})}; \quad \mathrm{sign} = -\,\mathrm{sign}\big(\Re[N_\mathbf{i}]\Im[H_{ij}]\big) \quad (3)$$

and for imaginary parent walkers:

$$p_s^{\Re}(\mathbf{j}|\mathbf{i}) = \frac{\delta\tau|\Im[H_{ij}]|}{p_{\mathrm{gen}}(\mathbf{j}|\mathbf{i})}; \quad \mathrm{sign} = -\,\mathrm{sign}\big(\Im[N_\mathbf{i}]\Im[H_{ij}]\big) \quad (4)$$

$$p_s^{\Im}(\mathbf{j}|\mathbf{i}) = \frac{\delta\tau|\Re[H_{ij}]|}{p_{\mathrm{gen}}(\mathbf{j}|\mathbf{i})}; \quad \mathrm{sign} = -\,\mathrm{sign}\big(\Im[N_\mathbf{i}]\Re[H_{ij}]\big) \quad (5)$$

where τ is the timestep for the simulation, and $p_s^{\Re}$ and $p_s^{\Im}$ indicate the probability of creating real and imaginary child walkers, respectively. $p_{\mathrm{gen}}(\mathbf{j}|\mathbf{i})$ is the probability of generating determinant $D_\mathbf{j}$ from $D_\mathbf{i}$. After this step, the 'death' step is performed for each occupied determinant, with the same death probability for the real and imaginary walkers of $\delta\tau(H_{ii} - E_0 - S)$ stochastically realized, where E_0 is a reference energy, and H_{ii} is now strictly real. A final annihilation step occurs every iteration, where real and imaginary walkers are separately considered, and pairs of opposite sign on the same determinant are removed from the simulation.

The value of S can be used as a strictly real measure of the correlation energy of the problem; however, provided a good overlap of the walker distribution with a reference wavefunction D_0 (generally taken to be the Hartree–Fock determinant) can be found, an averaged projected estimator is often less noisy:

$$E(\tau) = \frac{\langle D_0|\hat{H}|\Psi(\tau)\rangle}{\langle D_0|\Psi(\tau)\rangle} \quad (6)$$

where $\langle D_0|\Psi(\tau)\rangle = N_0$. As opposed to S, $E(\tau)$ is now a complex quantity, where in order to achieve real energies, the imaginary part of the energy must cancel to zero in a non-trivial way. In order to test this, we considered rock-salt-structured LiH sampled using $2\times2\times2$ k-points. By choosing all k-points to lie at the Γ-point or Brillouin zone boundary, it was possible to take linear combinations of the orbitals to give a strictly real basis. This is compared to the complex basis in Fig. 2.

It can be seen that all methods, including FCIQMC, agree exactly between the two bases, and that in the complex basis, the imaginary component of the energy converges to zero within small error bars. Although there is the potential for rotations of the wavefunction in the complex plane, it is observed that the discretization of the wavefunction amplitudes prevents this from happening, and global U(1) transformations are thus suppressed after an initial arbitrary phase factor is determined. This indicates that there is no more of a sign-issue to overcome with annihilation events than that of the original real formulation of the dynamics.

Real solids and the accuracy of quantum chemistry

Having established the accuracy and efficiency of the complex FCIQMC walker dynamics, it was initially tested on the most widely studied solid to date, rock-salt-structured LiH (refs 6–9, 11, 17, 27). We first benchmark the accuracy of the MP2, CCSD and CCSD(T) energies for this system, by considering the deviation of each from FCIQMC values. Figure 3 shows an equation of state for a range of volumes, with a $3\times3\times3$ k-point sampling, in a minimal basis required to capture any non-dynamic correlation.

The efficient sampling of the i-FCIQMC method is clear, where the $3\times3\times3$ k-point mesh correlates 54 electrons in a space of $\sim10^{30}$ determinants, with convergence to the exact energy obtained after
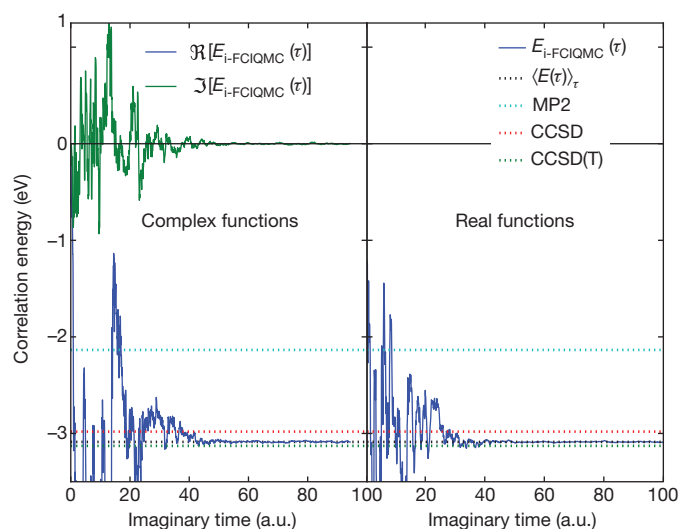

Figure 2 | Comparison of real and complex i-FCIQMC dynamics for the correlation energy of LiH. A $2\times2\times2$ Γ-centred k-point mesh with 16 electrons, and 40 correlated Hartree–Fock orbitals was employed, at a primitive rock-salt unit cell volume of $17.03\,\text{Å}^3$. Converged energies for 30 million walkers between the two bases agree within small stochastic error bars. The additional overhead for the complex dynamic means that the cost was ~5 times that of the real dynamic to converge to equivalent error bars. Also included are MP2, CCSD and CCSD(T) results for comparison. $\langle E(\tau)\rangle_\tau$ is an imaginary-time average of the projected energy $E(\tau)$, taken after a period of equilibration.

only ~50 million walkers, as demonstrated in Fig. 4a. The MP2 values are clearly shifted to higher energies compared to FCIQMC, and because this error changes significantly with volume, the MP2 equilibrium volume and bulk modulus deviate by 3.5% and 6.5%, respectively, from the FCIQMC values.

CCSD energies are virtually parallel to the FCIQMC results, yielding a similar volume and bulk modulus as FCIQMC. Finally, CCSD(T) shows almost exact agreement with FCIQMC in the absolute energies. Overall, these results very much mimic the performance of the standard quantum-chemical hierarchy established for molecular systems and are indicative of their suitability for other similar solids.
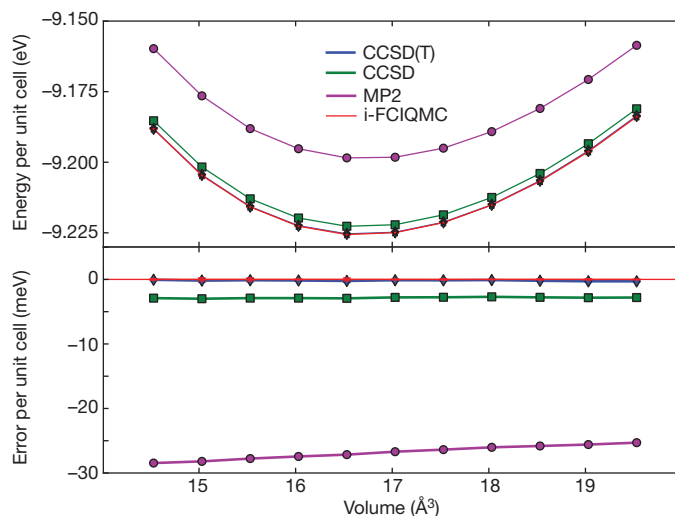

Figure 3 | Equation of state for rock-salt LiH, with the error from the i-FCIQMC value per unit cell at each volume shown below, for MP2, CCSD and CCSD(T). A $3\times3\times3$ Γ-centred k-point mesh was employed, with 54 electrons in 54 Hartree–Fock orbitals. All i-FCIQMC energies were converged with 55 million walkers, and error bars are too small to be seen on the plot ($\mathcal{O}[0.1\,\text{meV}]$).
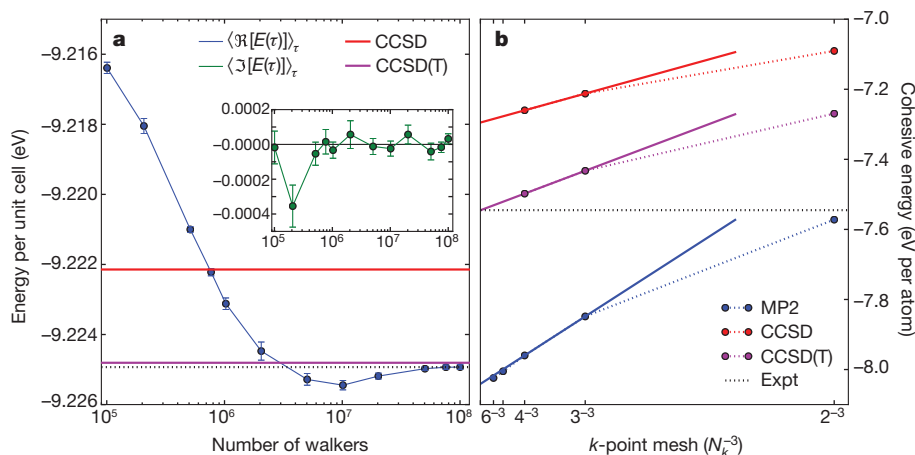
Figure 4 | Convergence with respect to the total walker number and *k*-point meshes. a, Convergence of the i-FCIQMC energy with total walker number of a $3 \times 3 \times 3$ *k*-point sampled LiH calculation. This explicitly correlated 54 electrons in 54 orbitals. Error bars (1s.d.) were calculated via a Flyvbjerg–Petersen 'blocking' algorithm[43]. The inset shows the imaginary component of the energy on a finer scale. **b**, Convergence of the cohesive energy of diamond with respect to the *k*-point mesh using MP2, CCSD and CCSD(T). The solid lines extrapolate to an infinite *k*-points mesh using $3 \times 3 \times 3$ and $4 \times 4 \times 4$ *k*-points. MP2 results for $5 \times 5 \times 5$ and $6 \times 6 \times 6$ *k*-points confirm that the finite size error decays as $1/N_k^3$, where N_k is the number of *k*-points in each direction.

To cover different bonding situations, in Fig. 5 we consider the relative errors of quantum-chemical methods for the correlation energies of several crystals when compared to FCIQMC. MP2 and CCSD recover between 80% and 98% of the FCIQMC correlation energy for the sample of rare gas, covalent and ionic solids. The dependence of the relative errors on the various systems is most pronounced in the case of MP2 theory. This is not unexpected and reflects the limitations of low-order perturbation theory. MP2 is more accurate for wide-gap insulators, such as Ne and LiF, than for semiconductors with a smaller gap, like Si and AlP. In contrast, CCSD(T) is shown to give a balanced description across the different systems and is in error by at most 2%.

What accuracy, therefore, can one expect from converged CCSD(T) calculations for solids? To answer this question, we computed cohesive energies. These are extremely demanding quantities for any theory, because the correlation effects in solids differ markedly from atoms, potentially leading to large errors in the prediction of the cohesive energy. Furthermore, attention needs to be paid to finite-size scaling (that is, *k*-point sampling). For the present study we have limited our attention to four solids, rock-salt LiH, diamond, zinc-blende BN and

AlP, expecting similar results for the other materials. We have used the progressive downsampling technique[14,28], employing *k*-point meshes of up to $4 \times 4 \times 4$. Figure 4b shows for diamond that the MP2 cohesive energy converges as $1/N_k^3$, where N_k is the number of *k*-points used to sample the Brillouin zone in each direction. By fitting to the MP2, CCSD and CCSD(T) energies for $3 \times 3 \times 3$ and $4 \times 4 \times 4$ *k*-points, we can extrapolate to infinitely dense *k*-point meshes. The remaining finite size error on the correlation energy is expected to be less than 20 meV per atom for the considered systems. The computational cost is of the order of 25,000 CPU (central processing unit) hours for diamond, with results reported in Table 1.

As anticipated (owing to the established agreement with i-FCIQMC in smaller bases and supercells), the CCSD(T) results are all in almost exact agreement with the experimental cohesive energy corrected for the zero-point energy[38]. The MP2 cohesive energies generally show substantial error compared to experiment[5,9,17]. MP2 severely underestimates the correlation energy of atoms, and while it also underestimates the correlation energy of the solids as shown in Fig. 5, the perturbative nature of the theory leads to less underestimation of the correlation energy in solids so that the cohesive energy is often—but not always—overestimated compared to the experimental value. This makes MP2 an unreliable method for solids, especially for calculation of cohesive energies. CCSD and CCSD(T), on the other hand, are far more consistent. Although the absolute errors in the CCSD cohesive energies are still quite sizeable, the cohesive energies are always underestimated. This results from the CCSD correlation energy in the solids being always too small (Fig. 5) whereas in atoms it is generally a good approximation. On adding the perturbative triples correction to the CCSD, the cohesive energies dramatically improve, to an error of only 0.03 eV. As for the absolute correlation energies in solids, the (T) correction over-compensates, leading to too-negative correlation energies, and hence an overestimation of the cohesive energies, albeit only mildly so. We finally note that our results indicate that the residual errors are dominated by correlation errors in the
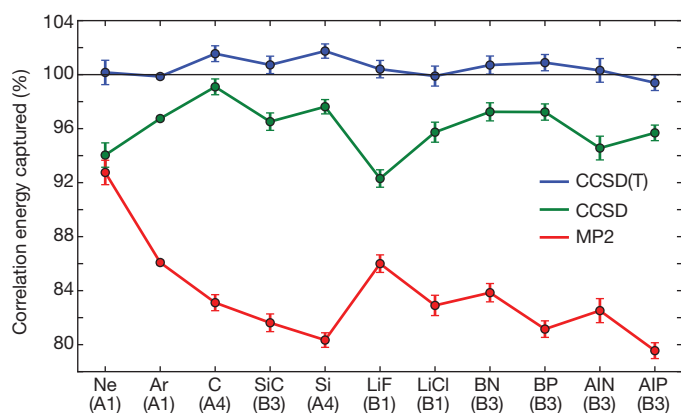

Figure 5 | Relative errors compared to i-FCIQMC in a range of solids. $2 \times 2 \times 2$ *k*-point sampling was employed, correlating 64 electrons in 64 orbitals, and 80 electrons in 72 orbitals, for LiF and LiCl, respectively. Lattice structures are given as A1 = f.c.c., A4 = diamond, B1 = rock salt, B3 = zinc blende, while lattice constants can be found in ref. 44. Error bars of each method are derived from the random errors of the i-FCIQMC values they are compared to. The i-FCIQMC calculations employed ~10,000 walkers on the Hartree–Fock determinant resulting in between 2 million and 300 million total walkers, depending on the system.

Table 1 | Cohesive energies evaluated within the quantum-chemical hierarchy

	LiH	C	BN	AlP
MP2	−2.386	−8.039	−7.149	−4.629
CCSD	−2.454	−7.295	−6.572	−4.107
CCSD(T)	−2.483	−7.545	−6.782	−4.347
Experiment	−2.487	−7.545	−6.758	−4.322

Energies are given in eV per atom. Experimental values have been corrected for zero-point vibration energies.
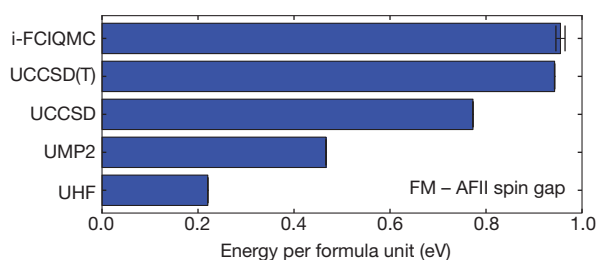
Figure 6 | Energies of the NiO spin gap, calculated at the level of UHF, UMP2, UCCSD, UCCSD(T) and i-FCIQMC in an UHF basis for both spin species. Correlating the $3d$ and $4s$ of the Ni, and $2s$ and $2p$ of the O using a rhombohedral supercell with a lattice constant of 4.19 Å containing 2 formula units, resulted in a sampling of 32 electrons, in 38 Hartree–Fock orbitals. 300 million walkers were required for the i-FCIQMC calculation of the AFII state, requiring ~50,000 CPU hours, while only 100 million walkers were needed for the FM state, with random error bar (1 s.d.) shown, and remaining error estimated to be less than 30 meV.

solid, whereas correlation energies for atoms are essentially converged at the CCSD(T) level.

The conclusion is that the accuracy of CCSD(T) has been established for solids to be of the order of 0.03 eV or 1 kcal mol^{-1}. By contrast, the most widely used density functional (PBE[39]) exhibits a mean-absolute error of 0.15–0.2 eV for a similar range of insulating solids[40].

Towards strong correlation

An important and open question is the domain of applicability of CCSD and CCSD(T) as stronger correlation effects set in. An initial indication of this can be found by comparing the performance of the methods for the spin gap between the ground antiferromagnetic state (AFII) and the ferromagnetic state (FM) of nickel oxide in a rhombohedral unit cell. This is a classic charge-transfer insulator, and it is expected to have strong correlation effects for at least one of the states, and as such, a balanced description for the calculation of the spin gap is expected to provide a stern test.

The results for the FM–AFII spin gap are given in Fig. 6 and illustrate the systematic convergence of the quantum-chemical hierarchy with respect to the correlation treatment of the system. Analysis of the FCIQMC wavefunction indicates that the ground state (AFII) is more strongly correlated and multiconfigurational, with a normalized Hartree–Fock weight of only 0.69, compared to 0.86 for the FM state. This lack of a dominant single reference leads to errors in MP2 of over 50%, while CCSD is still 19% in error compared to i-FCIQMC. Despite this, the qualitative behaviour of the quantum chemical hierarchy remains intact, with CCSD(T) providing an excellent approximation to the exact result. Results from unrestricted Hartree–Fock (UHF) calculations of NiO converged to the thermodynamic limit only capture 17% of the experimental spin gap between the two states[41], inferred from neutron scattering[42]. This agrees well with our restricted k-point sampled system, where the UHF captures 23% of the spin gap compared to FCIQMC results.

Conclusions and outlook

We have shown that FCI-quality correlation energies can be obtained for solid-state systems using an extension of the i-FCIQMC method to complex wavefunctions. We have demonstrated that the standard quantum-chemical hierarchy of increasingly accurate polynomially scaling methods holds for a range of materials, including rare gas, ionic and covalent solids and the charge-transfer insulator NiO. As explicitly shown by the cohesive energies of LiH, C, AlP and BN, CCSD(T) is very accurate for the solid state, surpassing 1 kcal mol^{-1} accuracy in reproducing experimental results. Considering the proven reliability of CCSD(T) for molecules, we expect a similar precision for insulators and semiconductors in general, with metals

possibly requiring further methodological improvements. In combination with recent developments to reduce the computational cost—for instance with explicit inclusion of the cusp condition for the many-electron wavefunction—as well as further technical, algorithmic and methodological advances, the accuracy of FCIQMC and the quantum-chemistry methods will be brought routinely to solid-state physics and computational materials science. We are witnessing a slow but steady change of our computational paradigm.

1. Kohn, W. Nobel lecture: Electronic structure of matter–wave functions and density functionals. *Rev. Mod. Phys.* **71,** 1253–1266 (1999).
2. Cohen, A. J., Mori-Sánchez, P. & Yang, W. Challenges for density functional theory. *Chem. Rev.* **112,** 289–320 (2012).
3. Pople, J. A. Nobel lecture: Quantum chemical models. *Rev. Mod. Phys.* **71,** 1267–1274 (1999).
4. Müller, C. & Paulus, B. Wavefunction-based electron correlation methods for solids. *Phys. Chem. Chem. Phys.* **14,** 7605–7614 (2012).
5. Maschio, L. *et al.* Fast local-MP2 method with density-fitting for crystals. I. Theory and algorithms. *Phys. Rev. B* **76,** 075101 (2007).
6. Casassa, S., Halo, M., Maschio, L., Roetti, C. & Pisani, C. Beyond a Hartree-Fock description of crystalline solids: the case of lithium hydride. *Theor. Chem. Acc.* **117,** 781–791 (2007).
7. Ayala, P., Kudin, K. & Scuseria, G. Atomic orbital Laplace-transformed second-order Møller-Plesset theory for periodic systems. *J. Chem. Phys.* **115,** 9698–9707 (2001).
8. Usvyat, D. *et al.* Approaching the theoretical limit in periodic local MP2 calculations with atomic-orbital basis sets: the case of LiH. *J. Chem. Phys.* **134,** 214105 (2011).
9. Marsman, M., Grüneis, A., Paier, J. & Kresse, G. Second-order Møller–Plesset perturbation theory applied to extended systems. I. Within the projector-augmented-wave formalism using a plane wave basis set. *J. Chem. Phys.* **130,** 184103 (2009).
10. Stoll, H. & Doll, K. Approaching the bulk limit with finite cluster calculations using local increments: the case of LiH. *J. Chem. Phys.* **136,** 074106 (2012).
11. Nolan, S. J., Gillan, M. J., Alfè, D., Allan, N. L. & Manby, F. R. Calculation of properties of crystalline lithium hydride using correlated wave function theory. *Phys. Rev. B* **80,** 165109 (2009).
12. Knowles, P. J. & Handy, N. C. A new determinant-based full configuration-interaction method. *Chem. Phys. Lett.* **111,** 315–321 (1984).
13. Larsen, H., Olsen, J., Jørgensen, P. & Christiansen, O. Full configuration interaction benchmarking of coupled-cluster models for the lowest singlet energy surfaces of N$_2$. *J. Chem. Phys.* **113,** 6677–6686 (2000).
14. Grüneis, A. *et al.* Natural orbitals for wave function based correlated calculations using a plane wave basis set. *J. Chem. Theory Comput.* **7,** 2780–2785 (2011).
15. Shiozaki, T. & Hirata, S. Communications: Explicitly correlated second-order Møller–Plesset perturbation method for extended systems. *J. Chem. Phys.* **132,** 151101 (2010).
16. Hättig, C., Klopper, W., Köhn, A. & Tew, D. P. Explicitly correlated electrons in molecules. *Chem. Rev.* **112,** 4–74 (2012).
17. Nolan, S. J., Bygrave, P. J., Allan, N. L. & Manby, F. R. Comparison of the incremental and hierarchical methods for crystalline neon. *J. Phys. Condens. Matter* **22,** 074201 (2010).
18. Kwee, H., Zhang, S. & Krakauer, H. Finite-size correction in many-body electronic structure calculations. *Phys. Rev. Lett.* **100,** 126404 (2008).
19. Manby, F. R., Stella, M., Goodpaster, J. D. & Miller, T. F. Ill. A simple, exact density-functional-theory embedding scheme. *J. Chem. Theory Comput.* **8,** 2564–2568 (2012).
20. Booth, G. H., Thom, A. J. W. & Alavi, A. Fermion Monte Carlo without fixed nodes: a game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **131,** 054106 (2009).
21. Booth, G. H. & Alavi, A. Approaching chemical accuracy using full configuration-interaction quantum Monte Carlo: a study of ionization potentials. *J. Chem. Phys.* **132,** 174104 (2010).
22. Cleland, D., Booth, G. H. & Alavi, A. Survival of the fittest: accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.* **132,** 041103 (2010).
23. Cleland, D. M., Booth, G. H. & Alavi, A. A study of electron affinities using the initiator approach to full configuration interaction quantum Monte Carlo. *J. Chem. Phys.* **134,** 024112 (2011).
24. Booth, G. H., Cleland, D., Thom, A. J. W. & Alavi, A. Breaking the carbon dimer: the challenges of multiple bond dissociation with full configuration interaction quantum Monte Carlo methods. *J. Chem. Phys.* **135,** 084104 (2011).
25. Shepherd, J. J., Booth, G., Grüneis, A. & Alavi, A. Full configuration interaction perspective on the homogeneous electron gas. *Phys. Rev. B* **85,** 081103 (2012).
26. López Ríos, P. *et al.* Inhomogeneous backflow transformations in quantum Monte Carlo calculations. *Phys. Rev. E* **74,** 066701 (2006).
27. Binnie, S. J. *et al.* Bulk and surface energetics of crystalline lithium hydride: benchmarks from quantum Monte Carlo and quantum chemistry. *Phys. Rev. B* **82,** 165431 (2010).
28. Ohnishi, Y.-y. & Hirata, S. Logarithm second-order many-body perturbation method for extended systems. *J. Chem. Phys.* **133,** 034106 (2010).

29. Nemec, N., Towler, M. D. & Needs, R. J. Benchmark all-electron ab initio quantum Monte Carlo calculations for small molecules. *J. Chem. Phys.* **132,** 034111 (2010).

30. Morales, M. A., McMinis, J., Clark, B. K., Kim, J. & Scuseria, G. E. Multideterminant wave functions in quantum Monte Carlo. *J. Chem. Theor. Comput* **8,** 2181–2188 (2012).

31. Al-Saidi, W. A., Zhang, S. & Krakauer, H. Bond breaking with auxiliary-field quantum Monte Carlo. *J. Chem. Phys.* **127,** 144101 (2007).

32. Purwanto, W., Krakauer, H. & Zhang, S. Pressure-induced diamond to β-tin transition in bulk silicon: a quantum Monte Carlo study. *Phys. Rev. B* **80,** 214116 (2009).

33. Møller, C. & Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **46,** 618–622 (1934).

34. Čížek, J. On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods. *J. Chem. Phys.* **45,** 4256–4266 (1966).

35. Raghavachari, K., Trucks, G. W., Pople, J. A. & Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **157,** 479–483 (1989).

36. Troyer, M. & Wiese, U.-J. Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Phys. Rev. Lett.* **94,** 170201 (2005).

37. Spencer, J. S., Blunt, N. S. & Foulkes, W. M. The sign problem and population dynamics in the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **136,** 054110 (2012).

38. Schimka, L., Harl, J. & Kresse, G. Improved hybrid functional for solids: the HSEsol functional. *J. Chem. Phys.* **134,** 024116 (2011).

39. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77,** 3865–3868 (1996).

40. Csonka, G. I. *et al.* Assessing the performance of recent density functionals for bulk solids. *Phys. Rev. B* **79,** 155107 (2009).

41. Cora, F. *et al.* in *Density Functional Theory in Inorganic Chemistry, Structure and Bonding* (eds McGrady, J. & Kaltsoyannis, N.) 171–232 (Springer, 2004).

42. Hutchings, M. & Samuelsen, E. Measurement of spin-wave dispersion in NiO by inelastic neutron-scattering and its relation to magnetic properties. *Phys. Rev. B* **6,** 3447–3461 (1972).

43. Flyvbjerg, H. & Petersen, H. G. Error estimates on averages of correlated data. *J. Chem. Phys.* **91,** 461–466 (1989).

44. Grüneis, A., Marsman, M. & Kresse, G. Second-order Møller–Plesset perturbation theory applied to extended systems. II. Structural and energetic properties. *J. Chem. Phys.* **133,** 074107 (2010).

Author Contributions G.H.B. and A.G. contributed equally to this work. G.H.B. and A.A. developed the FCIQMC method, G.H.B. wrote the computer code, and A.G. developed the quantum-chemical methods in VASP. G.K. led the VASP project, and A.A. led the FCIQMC project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.H.B. (ghb24@cam.ac.uk) or A.G. (andreas.grueneis@univie.ac.at).

Nickel and helium evidence for melt above the core–mantle boundary

Claude Herzberg[1], Paul D. Asimow[2], Dmitri A. Ionov[3,4], Chris Vidito[1], Matthew G. Jackson[5] & Dennis Geist[6]

High ^{3}He/^{4}He ratios in some basalts have generally been interpreted as originating in an incompletely degassed lower-mantle source[1–9]. This helium source may have been isolated at the core–mantle boundary region since Earth's accretion[4–6]. Alternatively, it may have taken part in whole-mantle convection and crust production over the age of the Earth[7–9]; if so, it is now either a primitive refugium at the core–mantle boundary[8] or is distributed throughout the lower mantle[7,9]. Here we constrain the problem using lavas from Baffin Island, West Greenland, the Ontong Java Plateau, Isla Gorgona and Fernandina (Galapagos). Olivine phenocryst compositions show that these lavas originated from a peridotite source that was about 20 per cent higher in nickel content than in the modern mid-ocean-ridge basalt source. Where data are available, these lavas also have high ^{3}He/^{4}He. We propose that a less-degassed nickel-rich source formed by core–mantle interaction during the crystallization of a melt-rich layer or basal magma ocean[5,6], and that this source continues to be sampled by mantle plumes. The spatial distribution of this source may be constrained by nickel partitioning experiments at the pressures of the core–mantle boundary.

Primitive mantle has been estimated to contain 1,960 p.p.m. nickel (Ni)[10], and this is similar to estimates for depleted peridotite[11], which makes up the average mid-ocean-ridge basalt (MORB) source. It is also similar to the Ni content of fertile peridotite (which has less basalt than primitive peridotite but more melt than depleted peridotite so more basalt can be extracted during partial melting) obtained from recent high-quality measurements[12] (Supplementary Fig. 1 and Supplementary section 1), which can be described by:

$$\text{Ni (p.p.m.)} = 68.6\,\text{MgO (weight per cent)} - 630 \qquad (1)$$

Equation (1) fits the Ni contents of 110 fertile and depleted peridotite compositions to 92 p.p.m. ($\pm 2\sigma$); we refer to this as normal Ni peridotite. High-precision olivine analyses from these fertile and depleted peridotites[12] yield 2,800–3,100 p.p.m. Ni and magnesium (Mg)-numbers of 89–92 (that is, 100MgO/(MgO + FeO) in mole per cent; Fig. 1a, Supplementary Table 1 and Supplementary Information section 2). These olivines from mantle peridotite are very similar to primitive olivine phenocryst compositions from East Pacific Rise (EPR) basalts (Fig. 1a). They are also very similar to model Ni contents[13] shown by the black line in Fig. 1, which is appropriate for olivine phenocrysts of primary magmas having 8%–38% MgO and for melting of both fertile and depleted upper-mantle peridotite. Model olivines are also similar to observed olivine phenocrysts in hot Archaean komatiites from Alexo[14,15] (Fig. 1b) and other occurrences of high-temperature magmas (Methods), demonstrating restricted compositions over a broad temperature range[13].

Olivines from the Palaeocene-epoch picrites in Baffin Island and West Greenland, on the other hand, have higher Ni contents (up to 3,800 p.p.m.) than those from MORBs, Archaean komatiites, and mantle peridotite (Fig. 1c). Those with high Mg-numbers are also higher in Ni than olivines that are expected to crystallize from any partial melt of normal mantle peridotite, as indicated by Ca, Mn and Fe/Mn (Fig. 2). Primary magma major-element compositions were estimated from modelling of whole-rock lava compositions[16], but sorting of olivine phenocrysts in lava flows compromises reliable Ni estimates with this method. However, the Ni contents and Mg-numbers of the olivine phenocrysts themselves are unaffected by crystal sorting. We estimated a typical Ni content of the primary magma that is required to crystallize olivines with the observed Ni and highest Mg-numbers. The high-Ni olivines require a fertile peridotite source having ~2,360 p.p.m. Ni (Methods Summary). This is about 20% higher than normal Ni peridotite with 1,960 p.p.m. Ni and is well outside the 2σ uncertainty bounds of equation (1). The range of olivine compositions at low Mg-numbers and Ni contents is a consequence of variable olivine and clinopyroxene fractionation from primary magmas (Fig. 1c and Supplementary Information section 4).

The excess-Ni problem is not confined to Baffin Island and West Greenland. Although high-precision Ni data for olivine in intraplate occurrences are limited, we find high Ni in olivines from the Ontong Java plateau, Gorgona komatiites, and the Fernandina volcano in the Galapagos islands (Fig. 3). In all cases a peridotite source provenance is indicated by Ca, Mn and Fe/Mn (refs 13, 14) in olivine (Fig. 2 and Supplementary Information). The high Ni in Gorgona olivines is clearly seen because the sample suite includes olivines with very high Mg-numbers, close to those expected for early crystals from the primary magmas. The high Ni in olivines from the other occurrences is less clear, owing to their lower Mg-numbers. For these we modelled the Ni contents of olivines that would crystallize from primary magmas and their liquid lines of descent. Using primary magma compositions from ref. 16 with the Ni contents adjusted to those expected in melts from normal fertile peridotite containing 1,960 p.p.m. Ni (Fig. 1), the Ni contents of olivines that crystallize along liquid lines of descent are always lower than those observed in olivines from the Ontong Java plateau or the Fernandina volcano, regardless of the proportions of crystallizing olivine and clinopyroxene. Rather, all successful solutions require Ni-rich fertile peridotite sources having Ni contents up to 2,360 p.p.m.

Mantle that has been depleted by prior melting events can be high in both MgO and NiO compared with more fertile sources (see equation (1)). However, the Ni contents of partial melts of highly depleted peridotite and fertile peridotites are indistinguishable at liquid MgO < 25 weight per cent[13] (Supplementary Fig. 1) and the olivines of primary melts shown in Fig. 1c should be generally applicable. Therefore, Ni-rich olivine phenocrysts cannot crystallize from melts of normal depleted peridotite.

Excess Ni in olivine has also been attributed to high temperatures and pressures of melting[17]. However, this is not consistent with olivines

[1]Department of Earth and Planetary Sciences, Rutgers University, 610 Taylor Road, Piscataway, New Jersey 08854-8066, USA. [2]Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA. [3]Département de Géologie, Université J. Monnet (member of PRES Université de Lyon), 23 rue P. Michelon, 42023 Saint-Etienne Cedex 2, France. [4]Laboratoire Magmas et Volcans, CNRS UMR 6524, 42023 Saint Etienne, France. [5]Department of Earth and Environment, Boston University, 685 Commonwealth Avenue, Boston, Massachusetts 02215, USA. [6]Department of Geological Sciences, University of Idaho 3022, Moscow, Idaho 83844, USA.
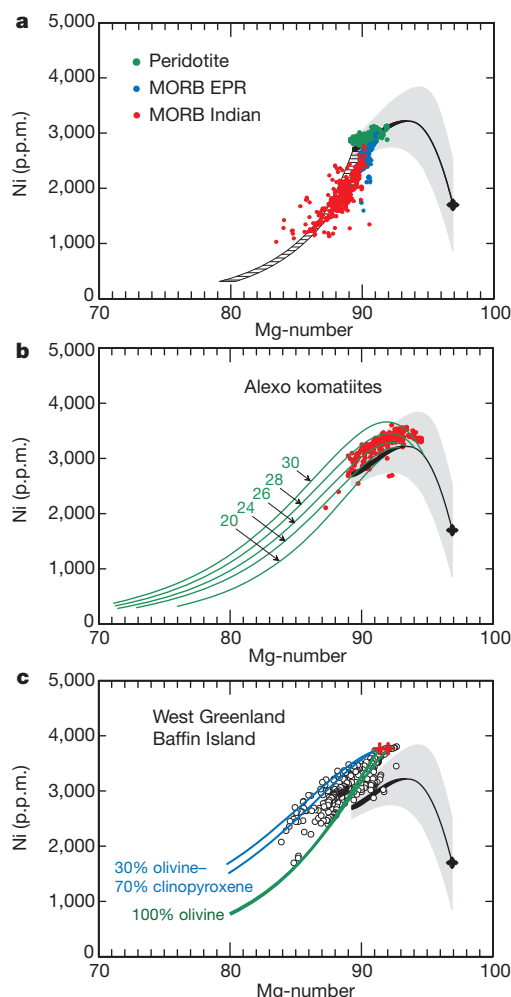
Figure 1 | Mg-numbers and Ni content for calculated[13] and observed olivine. An Mg-number is defined as 100MgO/(MgO + FeO)) in mole per cent. The black curve represents the calculated Ni content in olivines that crystallize from all primary melts derived from a fertile peridotite source having a whole-rock Ni content of 1,960 p.p.m. (ref. 13); the ±1σ uncertainty (grey shading) is discussed in Supplementary Information section 3. The black hatched area represents the calculated Ni content in olivines from olivine-fractionated derivative liquids from primary magmas having 8–13% MgO. **a**, The green circles are 203 high-precision analyses of olivine in mantle peridotite from various tectonic settings (Supplementary Information); the blue and red circles are high-precision olivine phenocrysts in MORBs from the Pacific (the EPR) and Indian oceans[14,17]. **b**, The red circles are high-precision olivine phenocrysts from the Archaean Alexo komatiites[14]. The green curves are calculated Ni contents in olivines from olivine-fractionated derivative liquids from primary magmas having 20–30% MgO (ref. 13), as indicated (green numbers); primary magmas for Alexo contained 26–30% MgO (refs 13, 15). **c**, The white circles are high-precision olivine phenocrysts from West Greenland and Baffin Island[14]; the uncertainty is only slightly greater than the size of the open circles. The red crosses are the Ni contents of olivine phenocrysts expected to crystallize from primary melts derived from Ni-rich fertile peridotite having 2,360 p.p.m. Ni. The green and blue lines are olivine compositions that will crystallize from these primary magmas by fractionation of olivine and clinopyroxene in the weight proportions indicated (Supplementary Information). A low-Ni population comes from one sample (ID 27142) from West Greenland[14], and it represents 28% of the database, having Mg-number > 90. There is little overlap between this sample and the more Ni-rich types. The difference between the low- and high-Ni populations most probably reflects mantle source heterogeneity, but there are no independent geochemical data for these samples with which to test this possibility.

from Archaean komatiites from Alexo[14] (Fig. 1b) and Barberton (see Methods), which are thought to be among the hottest magmas ever erupted on Earth[15], and yet have normal Ni concentrations. Appealing

to the formation of Alexo komatiites in a cool and wet environment does not solve the problem because they are compositionally similar to komatiites from the Belingwe greenstone belt[18] for which melt inclusion studies point to a hot and dry origin[19]. Furthermore, Barberton komatiites separated from a deep garnet-bearing residue that was hot and dry, and yet there is no excess Ni (Methods); low-temperature hydrous melting models are also not consistent with olivine phenocryst compositions (Methods). Consequently, the high Ni contents of the Baffin Island and West Greenland olivines are more plausibly related to composition rather than temperature-pressure effects.

Shield volcanoes from Hawaii also contain olivine phenocrysts with increased Ni (refs 14, 17), and there is general agreement that Hawaii is melting from a source that is hotter than the ambient mantle[16,17]. High Ni contents in Hawaiian olivines have been used to infer both pyroxenite melting in the source[13,14] and high temperatures and pressures of peridotite melting[17]. Comparably high mantle potential temperatures have been inferred for both Hawaii and West Greenland occurrences (1,500–1,600 °C; ref. 16) and both erupted through thick lithosphere (Supplementary Information). If increased temperature and pressure is the main mechanism for producing olivine phenocrysts with high Ni contents[17], there should be similar Ni, Mn, Fe/Mn and Ca contents in olivines from Hawaii and from Baffin Island and West Greenland. This is not observed (Supplementary Fig. 8). Hawaiian olivines are higher in Ni by ~1,000 p.p.m. and they have substantially lower calcium (Ca) and manganese (Mn) and higher Fe/Mn contents than those from West Greenland and Baffin Island. For the latter, a peridotite source provenance is clearly indicated (Fig. 2 and Supplementary Fig. 8). By contrast, low Ca and low Mn and high Fe/Mn contents for Hawaiian olivines are an expected consequence of pyroxenite source melting, owing to residuum retention of Ca in clinopyroxene and Mn in garnet (Supplementary Fig. 5). It is pyroxenite that contributes to high Ni in Hawaiian olivines[13,14], not temperature and pressure effects[17]. However, a possible role for mixed Ni-rich and Ni-poor peridotite with pyroxenite remains to be evaluated for Hawaii.

We conclude that high Ni coupled to peridotite-generated levels of Ca, Mn and Fe/Mn (Figs 1–3 and Supplementary Figs 6 and 7) point to Ni-rich peridotite sources for Baffin Island, West Greenland, Ontong Java plateau, Isla Gorgona and Fernandina in the Galapagos islands. By 'Ni-rich', we refer to Ni contents that are significantly higher than those described by equation (1), in contrast with normal-Ni peridotite sources that melt to make modern oceanic crust at mid-ocean ridges (Fig. 1a).

Current estimates of the Ni content of fertile mantle peridotite[10,11] are based on averages which can vary by about 300 p.p.m. (ref. 13); the Ni content of olivine in such peridotites obtained from open-access web sources is also large[17]. Although our estimated Ni content for Ni-rich peridotite is roughly within this range, we emphasize that this coincidence has no meaning because whole-rock and olivine Ni data are often compiled from sources of questionable accuracy[13], and there is much less uncertainty in our more recent work[12] (Fig. 1 and Supplementary Fig. 1). Random ±300 p.p.m. variations in the Ni content of peridotite can result in olivines with Ni as low as 2,200 p.p.m. (ref. 13). Such low Ni contents have never been reported from high-precision olivine analyses in unmetasomatized mantle peridotite (Fig. 1a) and, when found in olivine phenocrysts[14], other explanations are more plausible. For example, some MORB olivines can be low in Ni (Fig. 1a), but this is plausibly explained by sequestration of Ni into a residual sulphide phase[13], which can remain stable in the residue up to 15% melting[20].

Lavas exhibiting Ni excess are also associated with elevated [3]He/[4]He. Picrites from Baffin Island and West Greenland have [3]He/[4]He up to 50 times the atmospheric value (Ra) (ref. 21), and they also have primitive Nd and Pb isotopic compositions[3]. For the others, the maximum [3]He/[4]He (Ra) is 29 for Fernandina[2] and 18 for Gorgona[22]. No helium data are currently available for lavas from the Ontong Java plateau, but they also have primitive Nd and Pb isotopic ratios[8].
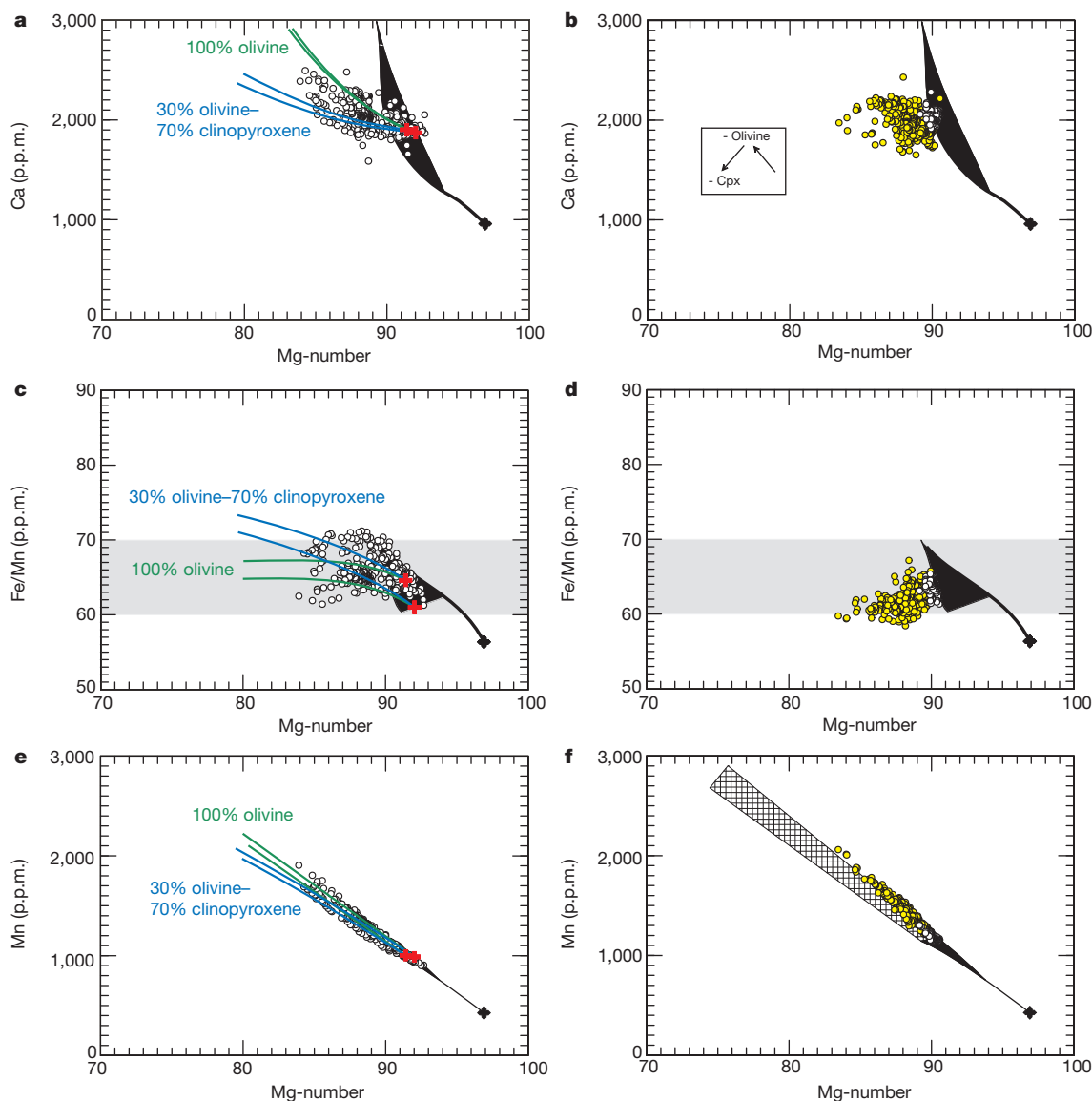
Figure 2 | Mg-numbers and Ca, Mn and Fe/Mn contents for calculated[13] and observed olivine phenocrysts from MORBs (right panels) compared with Baffin Island and West Greenland[14] (left panels). The black regions are calculated olivines from primary melts (8–38% MgO) derived by peridotite melting[13]. The black hatched area in **f** and the grey shading in **c** and **d** represent the calculated olivines from olivine-fractionated derivative liquids from primary magmas having 8–20% MgO. The green and blue lines in **a**, **c** and **e** are olivine compositions that crystallize from primary magmas that fractionated olivine and clinopyroxene in the weight proportions indicated (Supplementary Information). The inset in **b** shows how Ca in olivine changes owing to olivine and clinopyroxene fractionation from the primary magmas. The MORBs in **b**, **d** and **f** are from the EPR (white circles) and Indian Ocean (yellow circles). Although a wide range of Ca contents are theoretically possible for olivines that crystallize from peridotite-source primary magmas, most have ~2,000 p.p.m. Ca—similar to those found in MORBs, Baffin Island and West Greenland. Primitive and depleted peridotites contain ~1,000 p.p.m. Mn (refs 10, 11) and, although olivines with Fe/Mn = 55–70 can crystallize from primary magmas, more usually the range is restricted to 60–65 (ref. 13).

We now explore a model that might produce a Ni-rich peridotite with increased $^3He/^4He$. We have tested the formation of Ni-rich peridotite in a magma ocean by perovskite subtraction and/or ferropericlase addition. Although such models can yield high Ni contents, they produce a peridotite source that is too low in SiO_2 and Fe/Mn that is too high. More work on intra-mantle differentiation scenarios informed by accurate high-pressure phase equilibria and partitioning data are warranted[6], but here we propose a core–mantle interaction model.

Evidence for melting at the present-day core–mantle boundary includes the coincidence of recent determinations of the peridotite solidus with the estimated temperature at the top of the outer core[23] and seismic observations of ultralow-velocity zones[24]. As even more melt is expected in an early hotter Earth, present-day ultralow-velocity zones may represent the terminal stages in the crystallization of a long-lived magma ocean, and they provide a possible mechanism for

preserving primitive mantle[5,6]. This model predicts an equilibrium exchange of Ni between liquid metal in the core and liquid silicate in the mantle, which can be described by[25–27]:

$$FeO^{silicate} + Ni^{metal} = NiO^{silicate} + Fe^{metal} \qquad (2)$$

The present-day Ni content of fertile mantle peridotite (~1,960 p.p.m.) may have been established by core formation at ~30–60 GPa (ref. 25 and references therein). Silicate melt in contact with liquid metal would be enriched in Ni at 135 GPa as the exchange is driven from left to right with increasing pressure at constant oxidation state of iron[25–27]. We propose that a Ni-rich and less-degassed domain was formed by core–mantle interaction during the late stages of the crystallization of a basal magma ocean[5] or melt layer, and that some part of this source continues to be sampled by mantle plumes. Another
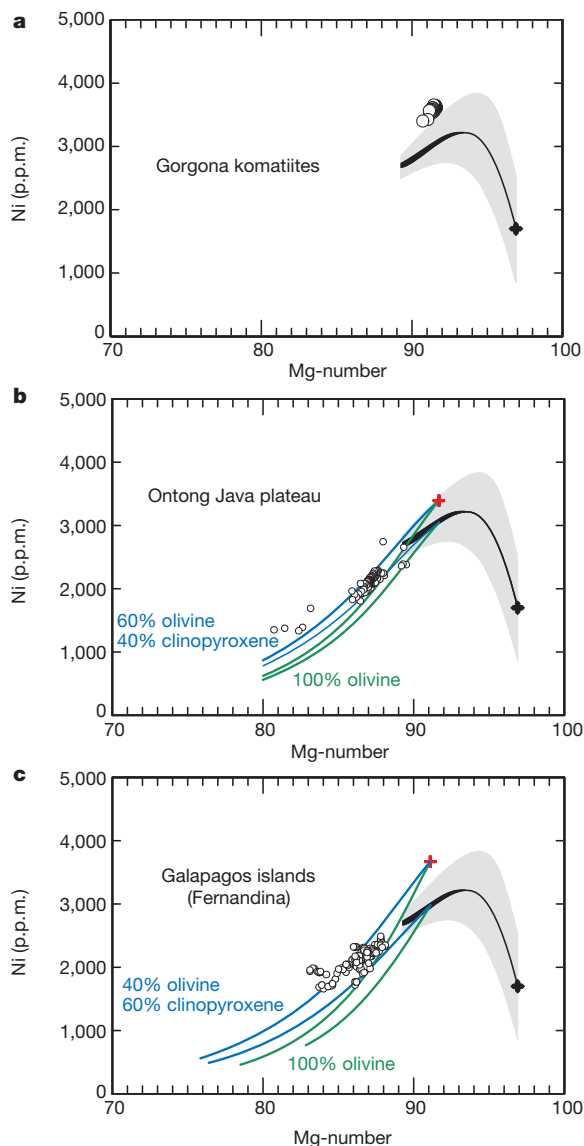
Figure 3 | Mg-numbers and Ni contents for calculated olivine[13] and observed olivine phenocrysts. a, Gorgona[14], **b**, the Ontong Java Plateau[14], and **c**, Fernandina (this work; Supplementary Information section 5 and Supplementary Table 2). The black curves represent the calculated Ni content in olivines of primary magmas of fertile peridotite having 1,960 p.p.m. Ni (ref. 13); the grey shaded regions (±1σ) represent a possible range of Ni contents in olivine that may arise from uncertainties in the calculation (Supplementary Information). The red crosses are olivines that would crystallize from Ni-rich fertile peridotite-source primary magmas; the green and blue lines are olivine compositions that crystallize from primary magmas that fractionated olivine and clinopyroxene in the weight proportions indicated (Supplementary Information section 4).

tracer for core–mantle interaction may be the addition of tungsten and osmium (Os) isotopes from the core to the mantle[27,28]. Gorgona is the only location examined here with available ^{186}Os/^{188}Os data, and there is an ^{186}Os anomaly that has been attributed to core influence[28]. There is no shortage of plausible mechanisms of core–mantle interaction, and the challenge is to understand how they are geochemically expressed[27].

Earth's mantle may have become oxidized by disproportionation of FeO in a deep magma ocean during crystallization of Fe^{3+}-rich Mg-perovskite and loss of metallic Fe to the core[29]. The Ni-rich peridotite source for West Greenland is also oxidized, having produced melts more oxidizing than MORB and within one log unit of the

nickel–nickel oxide oxygen fugacity buffer[30]. Although it is counter-intuitive, we propose that these oxidized magmas had an origin that began above the core–mantle boundary, in a location where Ni-rich silicate melt and Fe^{3+}-rich Mg-perovskite were in equilibrium with the outer core. Mantle plumes that sample this region and ascend above the melting zone would solidify to Ni-rich ferropericlase, Ca perovskite and Fe^{3+}-rich Mg-perovskite, transforming further to Ni-rich and oxidized peridotite in the upper mantle. Partial melting will yield magmas like those from West Greenland that are both Ni-rich and more oxidizing than MORB.

The spatial distribution of elevated ^{3}He/^{4}He in the mantle might be constrained by understanding the origin of a Baffin Island and West Greenland source with 2,360 p.p.m. Ni. Although core formation in a single stage is not likely[26], its prediction of 3,580 p.p.m. Ni (ref. 26) in the mantle at 135 GPa may be a guide to the Ni content of silicate melt at the present time. A source with 2,360 p.p.m. Ni would then require mixing of this high-Ni endmember with the canonical ~1,960 p.p.m. Ni from elsewhere in the mantle. In this scenario, the Ni-rich and high-^{3}He/^{4}He source may be widely distributed[7,9] or it may take the form of a halo around the core. Alternatively, the isolation of the Baffin Island and West Greenland source as a refuge from convection near the core–mantle boundary[4–6,8] is plausible if it can be demonstrated that 2,360 p.p.m. Ni in a silicate melt is in equilibrium with the core. In any case, interaction of the core with a silicate melt is a plausible mechanism for production of oxidized Ni-rich peridotite, and it provides evidence for a long-lived melt layer or basal magma ocean[5,6].

METHODS SUMMARY

The method for calculating model olivine Ni compositions is given elsewhere[13] and in the Supplementary Information. Primary magma compositions are calculated first, followed by the olivines that they crystallize.

The compositions of Ni in liquids extracted from dunite [liquid + olivine] and harzburgite [liquid + olivine + orthopyroxene] residues were computed by mass-balance solutions to the equation for accumulated fractional melting:

$$C_L = C_o[1 - (1 - F)^{1/D}]/F$$

where C_L is weight per cent NiO in the liquid (primary magma), C_o is the initial NiO in the peridotite source composition (0.25%, ref. 10; 1,960 p.p.m. Ni), F is the melt fraction, and D is the bulk distribution coefficient from the Beattie–Jones model (Supplementary Information). Ni contents for near-solidus melts were calculated from the MgO contents of near-solidus melts, together with a solidus olivine composition having 0.36% NiO (ref. 13) (Supplementary Information). Ni contents of near-solidus primary magmas are nearly indistinguishable from those of primary magmas for liquid + olivine and liquid + olivine + orthopyroxene assemblages, and the full range of possible Ni contents are shown in Supplementary Fig. 1.

The Ni content of olivine that crystallizes as a phenocryst at one atmosphere from a primary magma C_L was calculated using olivine/liquid Ni distribution coefficients from the Beattie–Jones model (Supplementary Information). Results are displayed in Figs 1 and 3 for a fertile peridotite having C_o = 1,960 p.p.m. Ni. Olivines have 2,800–3,100 p.p.m. Ni, and they provide an excellent description of olivine in normal mantle peridotite, primitive olivine phenocrysts in modern MORB, and olivine phenocrysts in Archaean komatiites. However, they fail to describe the much higher Ni contents of olivines from Baffin Island, West Greenland, the Ontong Java Plateau, Isla Gorgona and Fernandina (Galapagos) (Figs 1c and 3). For these, excellent agreement between computed and observed olivine Ni contents could be obtained with C_o = 0.30% NiO (2,360 p.p.m. Ni), a Ni-rich peridotite source.

1. Kurz, M. D., Jenkins, W. J. & Hart, S. R. Helium isotopic systematics of oceanic islands and mantle heterogeneity. *Nature* **297**, 43–47 (1982).

2. Kurz, M. D., Curtice, J., Fornari, D., Geist, D. & Moreira, M. Primitive neon from the center of the Galàpagos hotspot. *Earth Planet. Sci. Lett.* **286,** 23–34 (2009).
3. Jackson, M. G. *et al.* Evidence for the survival of the oldest terrestrial mantle reservoir. *Nature* **466,** 853–856 (2010).
4. Tolstikhin, I. & Hofmann, A. W. Early crust on top of the Earth's core. *Phys. Earth Planet. Inter.* **148,** 109–130 (2005).
5. Labrosse, S., Hernlund, J. W. & Coltice, N. A crystallizing dense magma ocean at the base of the Earth's mantle. *Nature* **450,** 866–869 (2007).
6. Coltice, N., Moreira, M., Hernlund, J. & Labrosse, S. Crystallization of a basal magma ocean recorded by helium and neon. *Earth Planet. Sci. Lett.* **308,** 193–199 (2011).
7. Class, C. & Goldstein, S. L. Evolution of helium isotopes in the Earth's mantle. *Nature* **436,** 1107–1112 (2005).
8. Jackson, M. G. & Carlson, R. W. An ancient recipe for flood-basalt genesis. *Nature* **476,** 316–319 (2011).
9. Gonnermann, H. M. & Mukhopadhyay, S. Preserving noble gases in a convecting mantle. *Nature* **459,** 560–563 (2009).
10. McDonough, W. F. & Sun, S.-s. The composition of the Earth. *Chem. Geol.* **120,** 223–253 (1995).
11. Salters, V. J. M. & Stracke, A. Composition of depleted mantle. *Geochem. Geophys. Geosyst.* **5,** Q05B07 (2004).
12. Ionov, D. A. Compositional variations and heterogeneity in fertile lithospheric mantle: peridotite xenoliths in basalts from Tariat, Mongolia. *Contrib. Mineral. Petrol.* **154,** 455–477 (2007).
13. Herzberg, C. Identification of source lithology in the Hawaiian and Canary islands: implications for origins. *J. Petrol.* **52,** 113–146 (2011).
14. Sobolev, A. V. *et al.* The amount of recycled crust in sources of mantle-derived melts. *Science* **316,** 412–417 (2007).
15. Arndt, N. T., Lesher, C. M. & Barnes, S. J. *Komatiite* 363–389 (Cambridge Univ. Press, 2008).
16. Herzberg, C. & Gazel, E. Petrological evidence for secular cooling in mantle plumes. *Nature* **436,** 619–622 (2009).
17. Putirka, K., Ryerson, F. J., Perfit, M. & Ridley, W. I. Mineralogy and composition of the oceanic mantle. *J. Petrol.* **52,** 279–313 (2011).
18. Puchtel, I. S., Walker, R. J., Brandon, A. D. & Nisbet, E. G. Pt–Re–Os and Sm–Nd isotope and HSE and REE systematics of the 2.7 Ga Belingwe and Abitibi komatiites. *Geochim. Cosmochim. Acta* **73,** 6367–6389 (2009).
19. Berry, A. J., Danyushevsky, D. V., O'Neill, H., St C., Newville, M. & Sutton, S. R. Oxidation state of iron in komatiitic melt inclusions indicates hot Archaean mantle. *Nature* **455,** 960–963 (2008).
20. Bézos, A., Lorand, J.-P., Humler, E. & Gros, M. Platinum-group element systematic in mid-oceanic ridge basaltic glasses from the Pacific, Atlantic, and Indian oceans. *Geochim. Cosmochim. Acta* **69,** 2613–2627 (2005).
21. Starkey, N. A. *et al.* Helium isotopes in early Iceland plume picrites: constraints on the composition of high ^{3}He/^{4}He mantle. *Earth Planet. Sci. Lett.* **277,** 91–100 (2009).
22. Révillon, S. *et al.* Heterogeneity of the Caribbean plateau mantle source: new constraints from Sr, O, and He isotope compositions of olivine and clinopyroxene. *Earth Planet. Sci. Lett.* **205,** 91–106 (2002).
23. Fiquet, G. *et al.* Melting of peridotite to 140 gigapascals. *Science* **329,** 1516–1518 (2010).
24. Williams, Q. & Garnero, E. J. Seismic evidence for partial melt at the base of Earth's mantle. *Science* **273,** 1528–1530 (1996).
25. Siebert, J., Badro, J., Antonangeli, D. & Ryerson, F. J. Metal-silicate partitioning of Ni and Co in a deep magma ocean. *Earth Planet. Sci. Lett.* **321/322,** 189–197 (2012).
26. Rubie, D. *et al.* Heterogeneous accretion, composition, and core-mantle differentiation of the Earth. *Earth Planet. Sci. Lett.* **301,** 31–42 (2011).
27. Walker, D. Core-mantle chemical issues. *Can. Mineral.* **43,** 1553–1564 (2005).
28. Brandon, A. D. *et al.* ^{186}Os-^{187}Os systematic of Gorgona Island komatiites: implications for early growth of the inner core. *Earth Planet. Sci. Lett.* **206,** 411–426 (2003).
29. Wade, J. & Wood, B. J. Core formation and the oxidation state of the Earth. *Earth Planet. Sci. Lett.* **236,** 78–95 (2005).
30. Larsen, L. M. & Pedersen, A. K. Processes in high-Mg, high T magmas: evidence from olivine, chromite and glass in Paleogene picrites from West Greenland. *J. Petrol.* **41,** 1071–1098 (2000).

Supplementary Information is available in the online version of the paper.

Author Contributions C.H. modelled olivine compositions and developed a variety of magma ocean and core–mantle interaction interpretations. P.D.A. suggested the core–mantle interaction model in its current form and critiqued all Ni partition models. D.A.I. provided high-precision olivine and whole-rock analyses for mantle peridotite. C.V. acquired high-precision olivine data for Fernandina (Galapagos). M.G.J. provided information on Pb, Nd and He isotopes. D.G. provided rock samples from Fernandina. All authors contributed to the intellectual growth of this paper.

METHODS

Our conclusion about a Ni-rich peridotite source is based on our use of the Beattie–Jones model[31,32] for the partitioning of Ni between olivine and liquid. We adopted it because it provides the minimum error when recovering experimental data compared with other published methods (Supplementary Information).

Putirka et al.[17] and Matzen et al.[33] argued that it is the temperature difference ΔT between separation from a high-pressure residue and onset of crystallization at shallower depths that accounts for the high Ni contents of olivine for Hawaii. This is termed a ΔT model. If this were true then it might be argued that this ΔT model is an alternative explanation to that of an Ni-rich peridotite source for Baffin Island, West Greenland, the Ontong Java Plateau, Isla Gorgona and Fernandina. However, the ΔT model of Putirka et al.[17] is based on a parameterization of experimental data with much larger errors than the Beattie–Jones Ni partitioning model (Supplementary Fig. 3 and Supplementary Information). We cannot comment on the ΔT model of Matzen et al.[33] because has not yet been published. Nevertheless, we examine in more detail the ΔT model because of its potential to compromise our interpretations.

Major-element, trace-element and isotope evidence suggest that the Barberton komatiites separated from a dry garnet-bearing peridotite residue at pressures in excess of 8 GPa (refs 34–38). The temperature difference between source and eruption conditions is about 300 °C for Barberton (see figure 5 in ref. 39), and about 150 °C for Hawaii[17,39]. The ΔT model predicts higher Ni contents of olivines in Barberton komatiites than in Hawaii. However, just the opposite is observed. Primitive Barberton olivines are lower in Ni than Hawaiian examples by more than 1,000 p.p.m. Ni (refs 36, 40) (Supplementary Figs 8 and 11; A. V. Sobolev, personal communication, 2012). This paradox might be resolved if the Barberton komatiites formed by low-temperature wet melting in a subduction environment[40–42] rather than dry hot melting in mantle plumes[34–39]. In this model melting is extensive: it goes beyond garnet stability, leaving behind a harzburgite residue. However, it is problematic because it ignores the aggregate body of evidence for residual garnet in heavy rare earth element depletions[36,38], CaO/Al$_2$O$_3$ and absolute major-element abundances[34,35], and positive ε_{Hf} together with subchondritic Lu/Hf[37].

The wet melting model is also not consistent with Barberton olivine compositions from the Komati Formation. Barberton olivines have Ca contents that are similar to olivines that crystallize from primary and derivative magmas of dry peridotite[36,40] (Supplementary Fig. 11). Previous calculations of Ca contents were based on accumulated fractional melting[13], but batch melting may have played a part at some stage owing to high melt densities at pressures above 8 GPa (refs 34, 36). Accordingly, olivines were computed from dry 7-GPa garnet-bearing melting experiments[35], and some are similar to olivines reported by ref. 40. The Ca contents of Barberton komatiites are consistent with melting that was hot, dry and deep, a conclusion that holds for either accumulated fractional melting or batch melting (Supplementary Fig. 11).

Experimental data show that the effect of magmatic H$_2$O is substantially to lower the Ca contents of olivines[43–45] (Supplementary Fig. 12), and it has been estimated that there was 5.4% to 7.1% magmatic H$_2$O for Barberton komatiite production[42,43]. If it is true that Barberton komatiites had ~6% magmatic H$_2$O

then this amount would be expected to drop the Ca content of olivine by ~1,000 p.p.m. relative to the dry system (Supplementary Fig. 12). Tenner et al.[45] drew attention to the possible relevance of their wet melting experiments on peridotite KLB-1 to the wet melting model for komatiites, and the MgO contents of their experimental melts (19%–27%) are roughly similar to the 23% MgO used in refs 40 and 41. However, a magmatic H$_2$O content of 6% is expected to crystallize olivines with ~1,000 p.p.m. Ca (Supplementary Fig. 12), in contrast to observed Barberton olivine phenocrysts, which average ~1,800 p.p.m. Ca (Supplementary Fig. 11). We conclude that the evidence for wet melting is weak and that Barberton komatiites separated from a deep garnet-bearing residue that was hot and dry. Normal Ni contents of Barberton olivines simply reflect the melting of a normal peridotite source. More complex models invoking temperature and pressure effects[17,33] are not consistent with olivine phenocryst compositions for Barberton komatiites.

31. Beattie, P. Ford, C. & Russell, D. Partition coefficients for olivine-melt and orthopyroxene-melt systems. *Contrib. Mineral. Petrol.* **109,** 212–224 (1991).

32. Jones, J. H. Temperature and pressure- independent correlations of olivine-liquid partition coefficients and their application to trace element partitioning. *Contrib. Mineral. Petrol.* **88,** 126–132 (1984).

33. Matzen, A. K., Baker, M. B., Beckett, J. R. & Stolper, E. M. Ni partitioning between olivine and high-MgO silicate melts: implications for Ni contents of forsteritic phenocrysts in basalts. *Abstr. Goldschmidt Conf.* http://www.vmgoldschmidt.org/2012/ (2012).

34. Herzberg, C. Geodynamic information in peridotite petrology. *J. Petrol.* **45,** 2507–2530 (2004).

35. Walter, M. J. Melting of garnet peridotite and the origin of komatiite and depleted lithosphere. *J. Petrol.* **39,** 29–60 (1998).

36. Robin-Popieul, C. C. M. et al. A new model for Barberton komatiites: deep critical melting with high melt retention. *J. Petrol.* **53 (11),** 2191–2229 (2012).

37. Blichert-Toft, J., Arndt, N. T. & Gruau, G. Hf isotopic measurements on Barberton komatiites: effects of incomplete sample dissolution and importance for primary and secondary signatures. *Chem. Geol.* **207,** 261–275 (2004).

38. Corgne, A. et al. Trace element partitioning between majoritic garnet and silicate melt at 10-17 GPa: implications for deep mantle processes. *Lithos* **148,** 128–141 (2012).

39. Herzberg, C., Condie, K. & Korenaga, J. Thermal history of the Earth and its petrological expression. *Earth Planet. Sci. Lett.* **292,** 79–88 (2010).

40. Parman, S. W., Dann, J. C., Grove, T. L. & de Wit, M. J. Emplacement conditions of komatiite magmas from the 3.49 Ga Komati Formation, Barberton Greenstone Belt, South Africa. *Earth Planet. Sci. Lett.* **150,** 303–323 (1997).

41. Grove, T. L. & Parman, S. W. Thermal evolution of the Earth as recorded by komatiites. *Earth Planet. Sci. Lett.* **219,** 173–187 (2004).

42. Grove, T. L., Parman, S. W. & Dann, J. C. in *Mantle Petrology: Field Observations and High Pressure Experimentation: A Tribute to Francis R. (Joe) Boyd* (eds Fei, Y., Bertka, C. M. & Mysen, B. O.) 155–167 (The Geochemical Society, Special Publication 6, 1999).

43. Parman, S. W. & Grove, T. L. Harzburgite melting with and without H2O: experimental data and predictive modeling. *J. Geophys. Res.* **109,** B02201 (2004).

44. Balta, J. B., Asimow, P. D. & Mosenfelder, J. L. Hydrous, low-carbon melting of garnet peridotite. *J. Petrol.* **52,** 2079–2105 (2011).

45. Tenner, T. J., Hirschmann, M. M. & Humayun, M. The effect of H2O on partial melting of garnet peridotite at 3.5 GPa. *Geochem. Geophys. Geosyst.* **13,** Q03016 (2012).

Planetary system disruption by Galactic perturbations to wide binary stars

Nathan A. Kaib[1,2]†, Sean N. Raymond[3,4] & Martin Duncan[1]

Nearly half the exoplanets found within binary star systems reside[1] in very wide binaries with average stellar separations greater than 1,000 astronomical units (one astronomical unit (AU) being the Earth–Sun distance), yet the influence of such distant binary companions on planetary evolution remains largely unstudied. Unlike their tighter counterparts, the stellar orbits of wide binaries continually change under the influence of the Milky Way's tidal field and impulses from other passing stars. Here we report numerical simulations demonstrating that the variable nature of wide binary star orbits dramatically reshapes the planetary systems they host, typically billions of years after formation. Contrary to previous understanding[2], wide binary companions may often strongly perturb planetary systems, triggering planetary ejections and increasing the orbital eccentricities of surviving planets. Although hitherto not recognized, orbits of giant exoplanets within wide binaries are statistically more eccentric than those around isolated stars. Both eccentricity distributions are well reproduced when we assume that isolated stars and wide binaries host similar planetary systems whose outermost giant planets are scattered beyond about 10 AU from their parent stars by early internal instabilities. Consequently, our results suggest that although wide binaries eventually remove the most distant planets from many planetary systems, most isolated giant exoplanet systems harbour additional distant, still undetected planets.

Unlike binary stars with separations below $\sim 10^3$ AU, very widely separated binaries ('wide binaries') are only weakly bound by self-gravity, leaving them susceptible to outside perturbations. As a result, the Milky Way's tide and impulses from other passing stars strongly perturb wide-binary orbits[3,4]. These perturbations, which are fairly independent of the orbiting object's mass, are also known to dramatically affect the dynamics of Solar System comets at similar orbital distances[5,6]. Galactic perturbations drive a pseudo-random walk in the pericentres of these comets (the closest approach distances to the Sun)[5,7]. The same effect will occur in wide-binary orbits. Thus, even if a very wide binary's initial pericentre is quite large, it will inevitably become very small at some point if it remains gravitationally bound and evolves long enough. Such low-pericentre phases will produce close stellar passages between binary members, with potentially devastating consequences for planetary systems in these binaries[8,9]. Counterintuitively, we therefore suspect that wide binary companions could more dramatically affect planetary system evolution than tight binaries.

To investigate this scenario, we use the Mercury simulation package to perform 2,600 simulations modelling the orbital evolution of our Sun's four giant planets (on their current orbits) in the presence of a very wide binary companion[10]. These simulations are listed as set A in Table 1, which briefly summarizes the initial conditions of our different simulations (see Supplementary Information for details). An example simulation is shown in Fig. 1. Initially, the binary companion has no effect on the planets' dynamics because its starting pericentre (q; solid black line in Fig. 1) is $\sim 3,000$ AU. However, after 1 Gyr of evolution, Galactic perturbations drive the binary pericentre near 100 AU, exciting the eccentricities of Neptune and Uranus. Once again, at 3.5 Gyr, the binary passes through another low-pericentre phase, this time triggering the ejection of Uranus. Last, at 7.2 Gyr, the binary makes a final excursion to low q, causing Neptune's ejection.

Such behaviour is not unusual. Depending on the binary's mass and semimajor axis (mean separation, or a_*), Fig. 2a demonstrates that ~ 30–60% of planetary systems in simulation set A experience instabilities causing one or more planetary ejections after 10 Gyr (the approximate age of our Galaxy's thin disk). Even though binaries with smaller semimajor axes are less affected by Galactic perturbations, Fig. 2a shows that the influence of binary semimajor axis on planetary instability rates is weak. This is because tighter binaries make pericentre passages at a higher frequency. In addition, when they reach low-q phases they remain stuck there for a much longer time than wider binaries. As Fig. 2b shows, both of these effects cause tighter binaries to become lethal (that is, destabilizing) at a much larger pericentre, offsetting the Galaxy's diminished influence. Most binary-triggered instabilities are very delayed. For binaries with $a_* \gtrsim 2,000$ AU, Fig. 2c shows that well over 90% of instabilities occur after at least 100 Myr of evolution, well after planet formation is complete. For tighter binaries, many more begin in orbits that destabilize the planets nearly instantly.

Although planets are believed to form on nearly circular orbits[11], most known giant planets ($m \sin i > 1\,M_{\mathrm{Jup}}$, where m is the planet's mass, i its orbital inclination and M_{Jup} is Jupiter's mass) have significant non-zero orbital eccentricities (eccentricities of less massive planets are known to be lower, or 'colder')[12]. This observed eccentricity distribution can be reproduced remarkably well when systems of circularly orbiting planets undergo internal dynamical instabilities causing planet–planet scattering events that eject some planets and excite the survivors' eccentricities[13–16]. For planetary systems within wide binaries, Fig. 2a predicts that many should undergo additional

Table 1 | Initial conditions of simulation sets

Name of simulation set	Number of planets	Planet masses (M_{Jup})	Planet a-range (AU)	Binary mass ($M_\odot$)	Binary $a*$ (AU)	External perturbations included
A	4	SS	SS	0.1–1.0	1,000–30,000	Tide + stars
B1	3	0.5 to ~15	2 to ~15	None	None	None
B2	3	0.5 to ~15	2 to ~15	0.4	1,000–30,000	Tide + stars
B3	3	0.5 to ~15	2 to ~15	0.4	1,000–30,000	None

SS refers to planetary systems resembling the Solar System's four giant planets. Tide + stars refers to perturbations from the Galactic tide and passing field stars. M_{Jup}, mass of Jupiter; a, semimajor axis; $M_\odot$, solar mass.

[1]Department of Physics, Queen's University, Kingston, Ontario K7L 3N6, Canada. [2]Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario M5S 3H8, Canada. [3]Université de Bordeaux, Observatoire Aquitain des Sciences de l'Univers, 2 rue de l'Observatoire, BP 89, F-33271 Floirac Cedex, France. [4]CNRS, UMR 5804, Laboratoire d'Astrophysique de Bordeaux, 2 rue de l'Observatoire, BP 89, F-33271 Floirac Cedex, France. †Present address: Center for Interdisciplinary Exploration and Research in Astrophysics and Department of Physics and Astronomy, Northwestern University, 2131 Tech Drive, Evanston, Illinois 60208, USA.
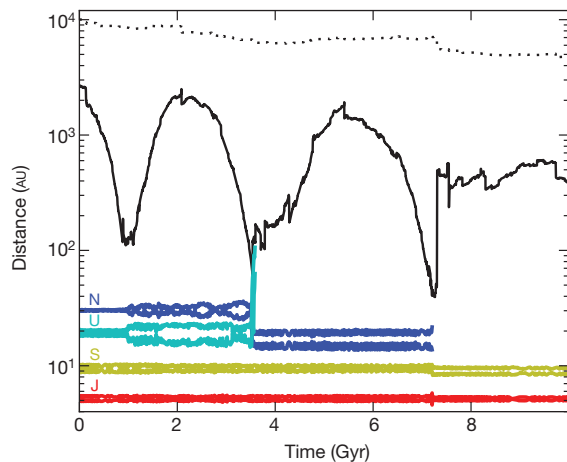
Figure 1 | Typical binary-triggered disruption. Simulation of a binary-triggered instability in a planetary system resembling our own Solar System. The pericentre and apocentre are plotted for Jupiter (J; red lines), Saturn (S; gold), Uranus (U; cyan) and Neptune (N; blue). The binary's semimajor axis (dotted black line) and pericentre (solid black line) are also shown.

dynamical instabilities triggered by their stellar companions. Thus, these systems should experience an even greater number of planet–planet scattering events than isolated planetary systems.

This raises the possibility that the eccentricities of exoplanets may hold a signature of the dynamical process illustrated in Fig. 1. Indeed, the overall distribution of exoplanet eccentricities provides compelling evidence of our disruptive mechanism. Figure 3a compares the observed eccentricity distribution of all Jovian-mass ($m \sin i > 1 M_{\mathrm{Jup}}$) exoplanets found in binaries[1] with the distribution of Jovian-mass planets around isolated stars. As can be seen, the distribution of planets within wide binaries is significantly hotter (or more biased to higher eccentricities) than planetary systems without known stellar companions. A Kolmogorov-Smirnov test returns a probability (P-value) of only 0.6% that such a poor match between the two data sets will occur if they sample the same underlying distribution. Thus, we reject the null hypothesis that the distributions are the same. Although it consists of just 20 planets, our wide-binary planetary sample contains the two most eccentric known exoplanet orbits, HD 80606b and HD 20782b (see Fig. 2). Furthermore, these excited eccentricities seem to be confined to only very wide binary systems. Figure 3a also shows the eccentricity distribution of planets residing in binaries with average separations below 10^3 AU. Unlike the case for wider binaries, here we see that these eccentricities match very closely with the isolated distribution. (A Kolmogorov-Smirnov test returns a P-value of 91%.) This suggests that the variable nature of distant binary orbits is crucial to exciting planetary orbits. Large eccentricities of planets within binaries have previously been explained with the Kozai resonance[17–20], yet this effect should be most evident in these tighter binary systems.

We perform additional simulations, attempting to explain the observed eccentricity excitation in Fig. 3a with the mechanism illustrated in Fig. 1. These additional simulation sets are summarized in Table 1 (B1–B3). Unlike the internally stable planetary systems in simulation set A, these simulated systems consist of three approximately Jovian-mass planets started in unstable configurations (to induce planet–planet scattering) and evolved for 10 Gyr (see Supplementary Information). In the simulation sets presented in Fig. 3b, we naturally reproduce both observed eccentricity distributions using the same initial planetary systems. When our planetary systems are run in isolation (set B1 in Table 1) planet–planet scattering caused by internal instabilities yields the observed planetary eccentricities for isolated stars (Kolmogorov-Smirnov test P-value of 42%). Then when a $0.4 M_\odot$ binary companion is added to each system (set B2 in Table 1) the eccentricity distribution is heated further, and again the match to
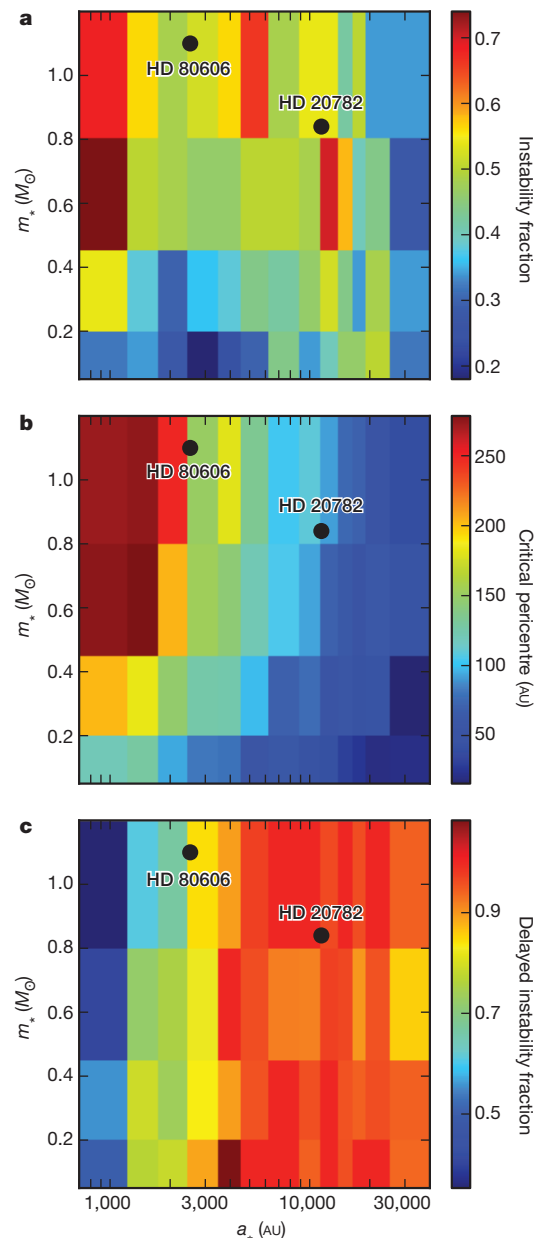

Figure 2 | Disruption as a function of binary mass and separation.
a–c, Maps of the fraction of systems in set A (see Table 1) that lost at least one planet via instability (**a**), the median binary pericentre below which an instability is induced in planetary systems (**b**) and the fraction of instabilities that occur after the first 100 Myr of evolution (**c**). In each panel, binary mass (m_*) is plotted on the y axis, and the binary semimajor axis (a_*) is plotted on the x axis; black data points mark the masses and presumed semimajor axes of the HD 80606 and HD 20782 binaries, which host the two most eccentric known planetary orbits[26,27]. Although HD 80606b has been reproduced with a Kozai-driven mechanism, this process is markedly slower in even wider binaries such as HD 20782b[28]. Moreover, the presence of more than one planet suppresses these Kozai oscillations[28–30]. However, our disruptive mechanism naturally collapses many systems to one planet, still enabling Kozai resonances to contribute to eccentricity excitation. Panel **a** suggests that binary-triggered instability rates become extremely high as binary semimajor axes drop below $\sim10^3$ AU, which could mean that tighter binaries trigger planetary system instabilities even more efficiently than those plotted here. However, the initial conditions assumed for both our planetary orbits (Solar System-like) and binary orbits (isotropic) become questionable for binary semimajor axes below $\sim10^3$ AU (see Supplementary Information). Another interesting aspect not immediately obvious in **c** is that instability times decrease at the largest binary semimajor axes. This is because such binaries are rapidly unbound (or 'ionized') by stellar impulses, making it impossible for these binaries to trigger instabilities at very late epochs (see Supplementary Information).
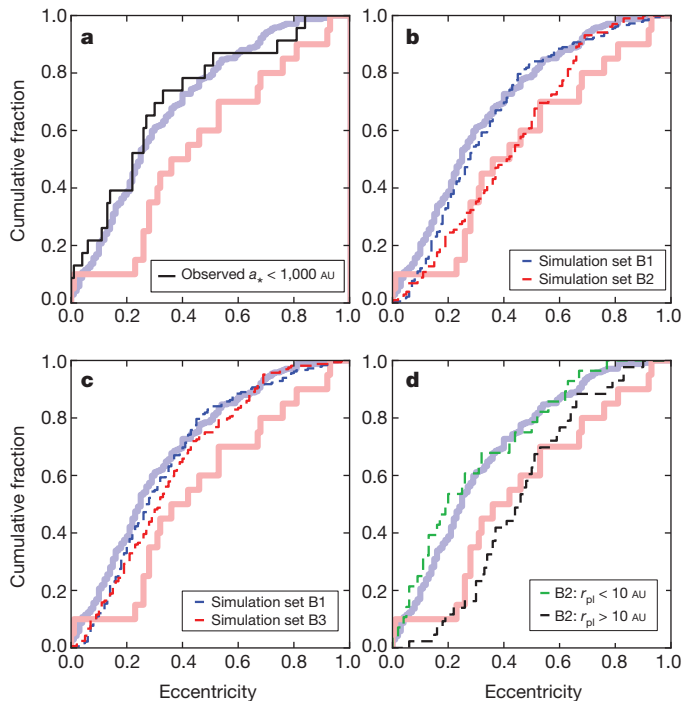
Figure 3 | Eccentricity excitation of planets of wide binaries. Comparison of planetary eccentricity distributions. In all panels, pale red lines correspond to observed systems with very wide ($a_* > 10^3$ AU) companions and pale blue lines are for observed isolated systems. Dashed lines describe the innermost planets in our simulated systems (sets B1–B3 in Table 1). In all distributions, planets with $a < 0.1$ AU are excluded to remove tidally circularized orbits. **a**, A comparison of observed exoplanet eccentricities within tighter ($a_* < 10^3$ AU) binaries (solid black line) to those observed in very wide ($a_* > 10^3$ AU) binaries and isolated systems. **b**, Eccentricities of simulated three-planet systems after 10 Gyr of evolution (sets B1 and B2 in Table 1). **c**, Simulations from **b** rerun with no Galactic perturbations (set B3 in Table 1). **d**, The final eccentricities of two different subgroups of wide binary simulations from **b**: systems that consisted of two planets extending beyond 10 AU at 10 Myr (black line), and two-planet systems confined inside 10 AU at 10 Myr (green). Notice in **b** that the presence of a wide binary does not seem to enhance the production of very extreme planetary eccentricities. However, one-quarter of our wide binary systems have planets driven into the central star (1.7 times the rate within isolated systems). Tidal dissipation not included in our models could strand these planets in very eccentric orbits before they collide with the central star[19,20]. Interestingly, binaries also completely strip 20% of our systems of planets, yielding naked stars that once hosted planets (see Supplementary Information).

observations is quite good, with a Kolmogorov-Smirnov test *P*-value of 46%.

In Fig. 3c, the match to observed planetary eccentricities is much poorer. Here we rerun our binary simulations with Galactic perturbations shut off to yield static binary orbits (set B3 in Table 1). In this case, the eccentricity distribution is barely more excited than the isolated cases, indicating that the variable nature of wide-binary orbits is crucial to heating planetary eccentricities. Otherwise, most stellar companions always remain far from the planets.

In Fig. 3d we reexamine simulation set B2 to determine which types of planetary systems are most influenced by wide binary companions. By examining the planetary systems after only 10 Myr, we can view them after most have experienced internal instabilities but before the binary has played a large role (because its effects are delayed). We find that 70% of our planetary systems have collapsed to two planets. (The remaining systems are composed of nearly equal numbers of one-planet and three-planet systems.) We then split these two-planet systems into those with the outer planet beyond 10 AU and those with both planets confined inside 10 AU. In Fig. 3d, the final ($t = 10$ Gyr) eccentricity distribution is shown for both subgroups of planetary systems. As can be seen, the more extended planetary systems eventually yield

much more excited eccentricities compared to the compact systems. This is because binaries do not have to evolve to such low pericentres to disrupt extended systems. In fact, the observed wide-binary planetary eccentricity distribution cannot be matched without using wide binaries with planets beyond 10 AU ($P = 1.6\%$ from a Kolmogorov-Smirnov test). Assuming planets form similarly in wide binaries and isolated systems, the planetary eccentricity excitation observed within wide binaries may offer new constraints on the bulk properties of isolated giant exoplanet systems, which dominate the giant exoplanet catalogue. Whereas most detection efforts are currently insensitive to planets with periods beyond ~10 yr, our work suggests that massive, longer-period planets (beyond ~10 AU) should be common around isolated stars. Indeed, such distant planets have recently been directly observed[21] and microlensing results suggest many such planets reside far from host stars[22].

Owing to the variable nature of their orbits, very distant binary companions may affect planetary evolution at least as strongly as their tighter counterparts. This represents a paradigm shift in our understanding of planet-hosting binaries, because previous work tended to assume that only tighter binaries strongly influence planetary system evolution[2,23]. Intriguingly, the eccentricities of planets in wide binaries may provide new constraints on the intrinsic architectures of all planetary systems. To develop this prospect further, searches for common proper motion companions to planet-hosting stars should be continued and expanded[2,23–25].

1. Roell, T., Neuhäuser, R., Seifahrt, A. & Mugrauer, M. Extrasolar planets in stellar multiple systems. *Astron. Astrophys.* **542,** A92 (2012).
2. Desidera, S. & Barbieri, M. Properties of planets in binary systems. The role of binary separation. *Astron. Astrophys.* **462,** 345–353 (2007).
3. Heggie, D. C. & Rasio, F. A. The effect of encounters on the eccentricity of binaries in clusters. *Mon. Not. R. Astron. Soc.* **282,** 1064–1084 (1996).
4. Jiang, Y.-F. & Tremaine, S. The evolution of wide binary stars. *Mon. Not. R. Astron. Soc.* **401,** 977–994 (2010).
5. Oort, J. H. The structure of the cloud of comets surrounding the Solar System and a hypothesis concerning its origin. *Bull. Astron. Inst. Neth.* **11,** 91–110 (1950).
6. Heisler, J. & Tremaine, S. The influence of the galactic tidal field on the Oort comet cloud. *Icarus* **65,** 13–26 (1986).
7. Kaib, N. A. & Quinn, T. Reassessing the source of long-period comets. *Science* **325,** 1234–1236 (2009).
8. Adams, F. C. & Laughlin, G. Constraints on the birth aggregate of the Solar System. *Icarus* **150,** 151–162 (2001).
9. Zakamska, N. L. & Tremaine, S. Excitation and propagation of eccentricity disturbances in planetary systems. *Astron. J.* **128,** 869–877 (2004).
10. Chambers, J. E., Quintana, E. V., Duncan, M. J. & Lissauer, J. J. Symplectic integrator algorithms for modeling planetary accretion in binary star systems. *Astron. J.* **123,** 2884–2894 (2002).
11. Lissauer, J. J. Planet formation. *Annu. Rev. Astron. Astrophys.* **31,** 129–172 (1993).
12. Wright, J. T. et al. Ten new and updated multiplanet systems and a survey of exoplanetary systems. *Astrophys. J.* **693,** 1084–1099 (2009).
13. Jurić, M. & Tremaine, S. Dynamical origin of extrasolar planet eccentricity distribution. *Astrophys. J.* **686,** 603–620 (2008).
14. Ford, E. B. & Rasio, F. A. Origins of eccentric extrasolar planets: testing the planet-planet scattering model. *Astrophys. J.* **686,** 621–636 (2008).
15. Malmberg, D. & Davies, M. B. On the origin of eccentricities among extrasolar planets. *Mon. Not. R. Astron. Soc.* **394,** L26–L30 (2009).
16. Raymond, S. N., Armitage, P. J. & Gorelick, N. Planet-planet scattering in planetesimal disks. II. Predictions for outer extrasolar planetary systems. *Astrophys. J.* **711,** 772–795 (2010).
17. Kozai, Y. Secular perturbations of asteroids with high inclination and eccentricity. *Astron. J.* **67,** 591–598 (1962).
18. Holman, M., Touma, J. & Tremaine, S. Chaotic variations in the eccentricity of the planet orbiting 16 Cygni B. *Nature* **386,** 254–256 (1997).
19. Wu, Y. & Murray, N. Planet migration and binary companions: the case of HD 80606b. *Astrophys. J.* **589,** 605–614 (2003).
20. Fabrycky, D. & Tremaine, S. Shrinking binary and planetary orbits by Kozai cycles with tidal friction. *Astrophys. J.* **669,** 1298–1315 (2007).
21. Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B. & Barman, T. Images of a fourth planet orbiting HR 8799. *Nature* **468,** 1080–1083 (2010).
22. The Microlensing Observations in Astrophysics (MOA) Collaboration & The Optical Gravitational Lensing Experiment (OGLE) Collaboration. Unbound or distant planetary mass population detected by gravitational microlensing. *Nature* **473,** 349–352 (2011).

23. Eggenberger, A. *et al.* The impact of stellar duplicity on planet occurrence and properties. I. Observational results of a VLT/NACO search for stellar companions to 130 nearby stars with and without planets. *Astron. Astrophys.* **474,** 273–291 (2007).

24. Raghavan, D. *et al.* Two suns in the sky: stellar multiplicity in exoplanet systems. *Astrophys. J.* **646,** 523–542 (2006).

25. Mugrauer, M. *et al.* A search for wide visual companions of exoplanet host stars: the Calar Alto survey. *Astron. Nachr.* **327,** 321–327 (2006).

26. Jones, H. R. A. *et al.* High-eccentricity planets from the Anglo-Australian Planet Search. *Mon. Not. R. Astron. Soc.* **369,** 249–256 (2006).

27. Naef, D. *et al.* HD 80606 b, a planet on an extremely elongated orbit. *Astron. Astrophys.* **375,** L27–L30 (2001).

28. Innanen, K. A., Zheng, J. Q., Mikkola, S. & Valtonen, M. J. The Kozai mechanism and the stability of planetary orbits in binary star systems. *Astron. J.* **113,** 1915–1919 (1997).

29. Batygin, K., Morbidelli, A. & Tsiganis, K. Formation and evolution of planetary systems in presence of highly inclined stellar perturbers. *Astron. Astrophys.* **533,** A7 (2011).

30. Kaib, N. A., Raymond, S. N. & Duncan, M. J. 55 Cancri: a coplanar planetary system that is likely misaligned with its star. *Astrophys. J.* **742,** L24 (2011).

Author Contributions N.A.K. performed the simulations and analysis and was the primary writer of this paper. S.N.R. and M.D. helped initiate the project and advised on simulations and analysis.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.A.K. (nkaib@astro.queensu.ca).

Exaggerated translation causes synaptic and behavioural aberrations associated with autism

Emanuela Santini[1], Thu N. Huynh[1], Andrew F. MacAskill[1], Adam G. Carter[1], Philippe Pierre[2,3,4], Davide Ruggero[5]*, Hanoch Kaphzan[1]†* & Eric Klann[1]

Autism spectrum disorders (ASDs) are an early onset, heterogeneous group of heritable neuropsychiatric disorders with symptoms that include deficits in social interaction skills, impaired communication abilities, and ritualistic-like repetitive behaviours[1,2]. One of the hypotheses for a common molecular mechanism underlying ASDs is altered translational control resulting in exaggerated protein synthesis[3]. Genetic variants in chromosome 4q, which contains the *EIF4E* locus, have been described in patients with autism[4,5]. Importantly, a rare single nucleotide polymorphism has been identified in autism that is associated with increased promoter activity in the *EIF4E* gene[6]. Here we show that genetically increasing the levels of eukaryotic translation initiation factor 4E (eIF4E) in mice[7] results in exaggerated cap-dependent translation and aberrant behaviours reminiscent of autism, including repetitive and perseverative behaviours and social interaction deficits. Moreover, these autistic-like behaviours are accompanied by synaptic pathophysiology in the medial prefrontal cortex, striatum and hippocampus. The autistic-like behaviours displayed by the eIF4E-transgenic mice are corrected by intracerebroventricular infusions of the cap-dependent translation inhibitor 4EGI-1. Our findings demonstrate a causal relationship between exaggerated cap-dependent translation, synaptic dysfunction and aberrant behaviours associated with autism.

eIF4E-transgenic mice (βT-*Eif4e*)[7] exhibited increased levels of eIF4E across brain regions (Fig. 1a) without compensatory changes in levels of other translational control proteins (Fig. 1b). We investigated whether eIF4E was bound preferentially to either eIF4E-binding protein (4E-BP) or eIF4G, which repress and promote, respectively, the initiation of cap-dependent translation[8,9]. We found significantly higher levels of eIF4E–eIF4G interactions in the brains of eIF4E-transgenic mice (Fig. 1c and Supplementary Fig. 1a), with no alterations in the interaction between eIF4E and 4E-BP (Fig. 1c, left, and Supplementary Fig. 1a). To confirm that the increased eIF4E–eIF4G interactions resulted in increased protein synthesis, we infused puromycin into the lateral ventricle of cannulated mice and labelled newly synthesized proteins using SUnSET[10,11], and observed increased *de novo* cap-dependent translation (Fig. 1d and Supplementary Fig. 1b–g). Overall, our results indicate that overexpression of eIF4E results in exaggerated cap-dependent translation in the brains of eIF4E-transgenic mice.

We then determined whether eIF4E-transgenic mice display repetitive and perseverative behaviours, which are behavioural domains required for ASD diagnosis[2]. eIF4E-transgenic mice exhibited repetitive digging behaviour in the marble-burying test[12] and increased self-grooming[13] compared with wild-type littermate controls (Fig. 2a, b). eIF4E-transgenic mice also displayed cognitive inflexibility in both a water-based Y-maze task and a modified version of the Morris water maze[14,15]. Learning ability in the acquisition and memory phases of

these tasks was intact; however, in the reversal phases, eIF4E-transgenic mice were impaired in locating the new platform positions (Fig. 2c, d and Supplementary Fig. 2e–h). We tested an additional form of behavioural inflexibility by examining the eIF4E-transgenic mice for extinction of cued fear conditioning and found that they did not exhibit a significant reduction in freezing responses after extinction training (Fig. 2e). These experiments suggest that excessive cap-dependent translation in the brain affects the ability to suppress previously codified response patterns and the ability to form new behavioural strategies in response to changed environmental circumstances.

Abnormalities in social interaction skills are another behavioural defect displayed by individuals with ASDs[2]. In tests to examine social behaviour[16–18], the eIF4E-transgenic mice did not show a preference for a nonspecific stranger versus a new, inanimate object (Fig. 2f, g). Moreover, eIF4E-transgenic mice exhibited diminished reciprocal interactions with a freely moving stranger mouse (Fig. 2h), further evidence of deficits in social behaviour. The deficits in social behaviour of the eIF4E-transgenic mice are unlikely to be caused by a generalized increase in anxiety (Supplementary Fig. 2c, d, j). Moreover, the eIF4E-transgenic mice exhibited mild hyperactivity (Supplementary Fig. 2a, b), but no impairments in motor coordination, motor learning and sensorimotor gating (Supplementary Fig. 2i, k, l). Taken together, our behavioural analysis of the eIF4E-transgenic mice indicates that increased cap-dependent translation in the brain results in a distinct pattern of behavioural abnormalities consistent with ASDs.

Previous studies suggest that ASD symptoms such as cognitive inflexibility and deficits in social behaviour are generated by abnormalities in prefrontal and/or striatal circuits[19]. Consistent with this idea, the medial prefrontal cortex (PFC) is implicated in the modulation of social behaviours and social skills[20], whereas motor, social and communication impairments in boys with ASDs are associated with anatomical abnormalities in the striatum[21]. Therefore, we next examined whether the eIF4E-transgenic mice exhibited specific synaptic pathophysiologies in the medial PFC and striatum.

In the eIF4E-transgenic mice, examination of spontaneous synaptic 'mini' events in layers 2/3 of acute medial PFC slices revealed an increase in the frequency but not amplitude of excitatory events (miniature excitatory postsynaptic currents (mEPSCs); Fig. 3a), and an increase in the amplitude, but not frequency, of inhibitory events (miniature inhibitory postsynaptic currents (mIPSCs); Fig. 3b). No changes were observed in layer 5 (Supplementary Fig. 3a, b). Thus, our data suggest an enhancement of excitatory input and postsynaptic sensitivity for inhibitory events onto layer 2/3 pyramidal neurons, consistent with the hypothesis that autism may arise from an imbalance between excitatory and inhibitory synaptic transmission[22].

To determine whether the increased frequency of spontaneous mEPSCs might result from an enhanced number of synaptic contacts, we imaged dendritic spines using two-photon laser-scanning
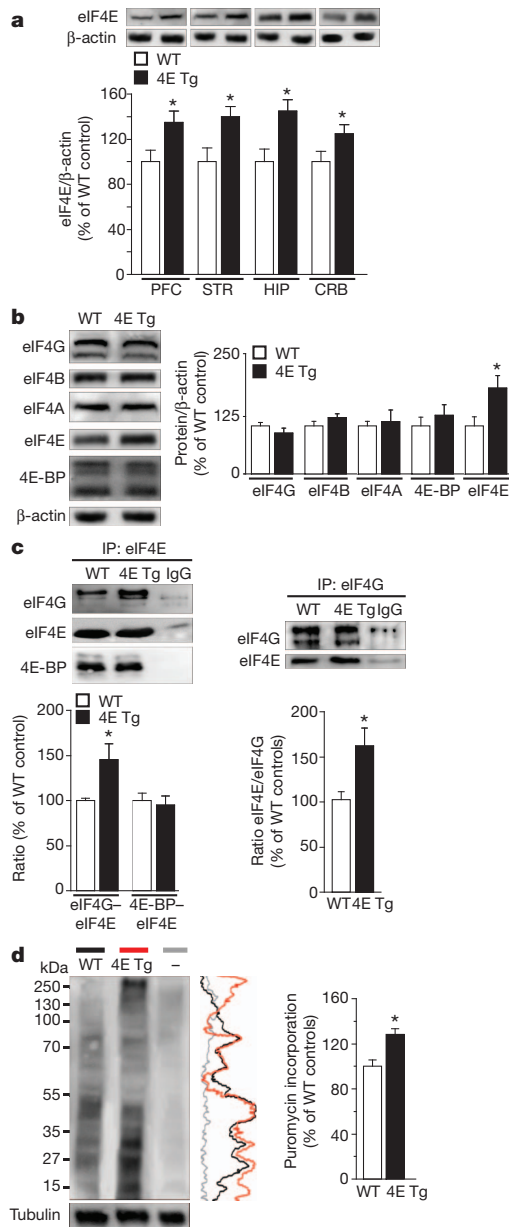
Figure 1 | eIF4E-transgenic mice exhibit increased eIF4E/eIF4G interactions and exaggerated cap-dependent translation. a, eIF4E-transgenic mice (4E Tg) exhibit increased eIF4E expression in multiple brain regions. $n = 4$ mice per genotype; $*P < 0.05$ versus wild type (WT), Student's t-test. CRB, cerebellum; HIP, hippocampus; PFC, prefrontal cortex; STR, striatum. **b,** eIF4E-transgenic mice exhibit normal expression of other translational control proteins. $n = 4$ mice per genotype; $*P < 0.05$, Student's t-test. **c,** eIF4E-transgenic mice exhibit increased eIF4E–eIF4G interactions. Immunoprecipitation (IP) of eIF4E (left) and eIF4G (right). $n = 3$ mice per genotype; $*P < 0.05$, Student's t-test. **d,** eIF4E-transgenic mice exhibit increased translation as measured by SUnSET (see Methods). Vertical line traces of each autoradiogram are shown on the right. $n = 3$ mice per genotype; $*P < 0.05$, Student's t-test. '–' represents a control sample without puromycin. All data are shown as mean and s.e.m.

Figure 2 | eIF4E-transgenic mice exhibit ASD-like behaviours. eIF4E-transgenic mice were compared to wild-type littermates. **a,** Marble-burying test. $n = 21$–22 mice per genotype; $*P < 0.05$, $***P < 0.001$, repeated-measures analysis of variance (ANOVA) (time × genotype, $F_{(2,46)} = 31.62$, $P < 0.001$) followed by Bonferroni–Dunn post-hoc test. **b,** Self-grooming test. $n = 12$ mice per genotype; $*P < 0.05$, Student's t-test. **c,** Y-maze reversal task. $n = 21$–22 mice per genotype; $*P < 0.05$, $***P < 0.001$, repeated-measures ANOVA (time × genotype, $F_{(5,138)} = 16.74$, $P < 0.001$) followed by Bonferroni–Dunn post-hoc test. **d,** Morris water maze reversal learning. $n = 12$–13 mice per genotype; $*P < 0.05$, repeated-measures ANOVA (time × genotype, $F_{(3,92)} = 6.1$, $P < 0.001$) followed by Bonferroni–Dunn post-hoc test. **e,** Extinction of cued fear memory (15 conditioned stimuli (CS) per day represented as three CS blocks). $n = 6$ mice per genotype; $*P < 0.05$, repeated-measures ANOVA (day 1: time × genotype, $F_{(4,40)} = 5.73$, $P < 0.001$; day 2: time × genotype, $F_{(4,40)} = 4.81$, $P < 0.01$) followed by Bonferroni–Dunn post-hoc test. **f, g,** Social behaviour test. The time spent either interacting with a stranger mouse (**f**) or in the chambers (**g**). $n = 6$ mice per genotype; $*P < 0.05$, repeated-measures ANOVA (stimulus × genotype, $F_{(1,10)} = 6.04$, $P < 0.05$ (**f**); stimulus × genotype, $F_{(1,10)} = 6.12$, $P < 0.05$ (**g**)) followed by Bonferroni–Dunn post-hoc test. **h,** Reciprocal social interaction task. $n = 6$ mice per genotype; $*P < 0.05$, Student's t-test. All data are shown as mean and s.e.m.

microscopy (Fig. 3c, d and Supplementary Fig. 3c, d). We found a significant increase (~12%) in spine density and observed a significantly smaller spine volume in the eIF4E-transgenic mice than in wild-type littermates (wild type = 0.123 ± 0.004 μm^3 (mean ± s.e.m.) and eIF4E-transgenic = 0.110 ± 0.004 μm^3, $P = 0.01$ versus wild type, Student's t-test).

Next, we examined whether increased expression of eIF4E also resulted in synaptic pathophysiology in the striatum. We used

high-frequency stimulation to induce long-term depression (LTD) in acute striatal slices[23], and found that eIF4E-transgenic mice exhibited enhanced LTD compared to wild-type littermates (Fig. 3e and Supplementary Fig. 3e, f). We propose that the enhanced LTD in eIF4E-transgenic mice results in altered efficiency of striatal information storage and processing, culminating in the inability to form new motor patterns and/or to disengage from previously learned motor behaviours.
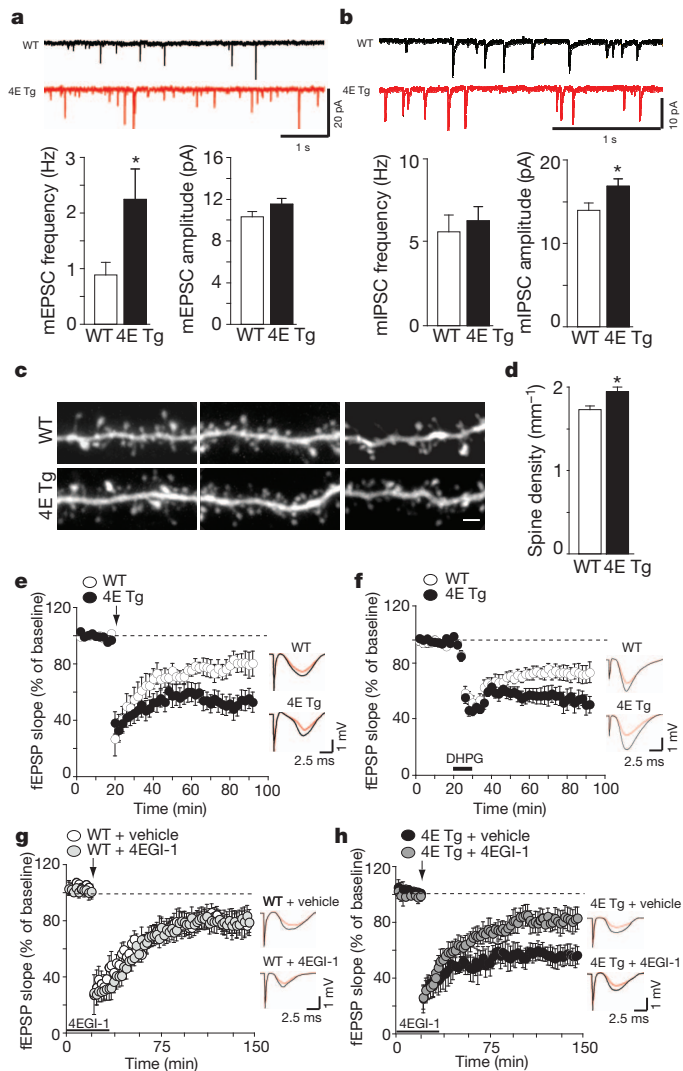
Figure 3 | eIF4E-transgenic mice exhibit alterations in synaptic function, dendritic spine density and synaptic plasticity. a, **b**, eIF4E-transgenic mice exhibit increased mEPSC frequency (**a**) and increased mIPSC amplitude (**b**) in layer 2/3 medial PFC pyramidal neurons. $n = 27$–30 neurons per genotype; *$P < 0.05$, Student's t-test. **c**, **d**, eIF4E-transgenic mice exhibit increased dendritic spine density in layer 2/3 medial PFC pyramidal neurons. High-magnification images (**c**) and quantification (**d**) of spiny dendrites. $n = 12$ neurons per genotype; *$P < 0.05$, Student's t-test. Scale bar, 2 μm. **e**, eIF4E-transgenic mice exhibit enhanced striatal LTD. $n = 13$ slices from 8 mice per genotype. **f**, eIF4E-transgenic mice exhibit enhanced hippocampal mGluR-LTD. DHPG denotes the mGluR agonist 3,5-dihydroxyphenylglycine. $n = 15$ slices from 8 mice per genotype. **g**, **h**, 4EGI-1 normalizes enhanced striatal LTD shown by eIF4E-transgenic mice (**h**), without affecting LTD in wild-type mice (**g**). $n = 18$ slices from 9 mice per genotype and treatment. All field recordings were analysed with repeated-measures ANOVA. Arrows indicate delivery of high-frequency stimulation. Solid bars indicate the duration of bath application of DHPG (10 μM, 10 min) and 4EGI-1 (100 μM, 45 min). Representative traces (right) showing field excitatory postsynaptic potentials (fEPSPs) before (black) and 60 min after (red) high-frequency stimulation. All data are shown as mean and s.e.m.

To determine whether the synaptic alterations described in the eIF4E-transgenic mice were selective for the frontostriatal circuit, we examined synaptic plasticity in the hippocampus[24]. We found that eIF4E-transgenic mice exhibited enhanced metabotropic glutamate receptor-dependent LTD (mGluR-LTD) compared to wild-type littermates (Fig. 3f and Supplementary Fig. 3g, h), consistent with previous studies showing that changes in brain protein synthesis are accompanied by altered (enhanced or reduced) hippocampal

mGluR-LTD[25,26]. Thus, consistent with the ubiquitous increase in brain expression of eIF4E, the eIF4E-transgenic mice display altered synaptic function and plasticity in several brain regions (medial PFC, striatum and hippocampus) implicated in behavioural abnormalities associated with ASDs.
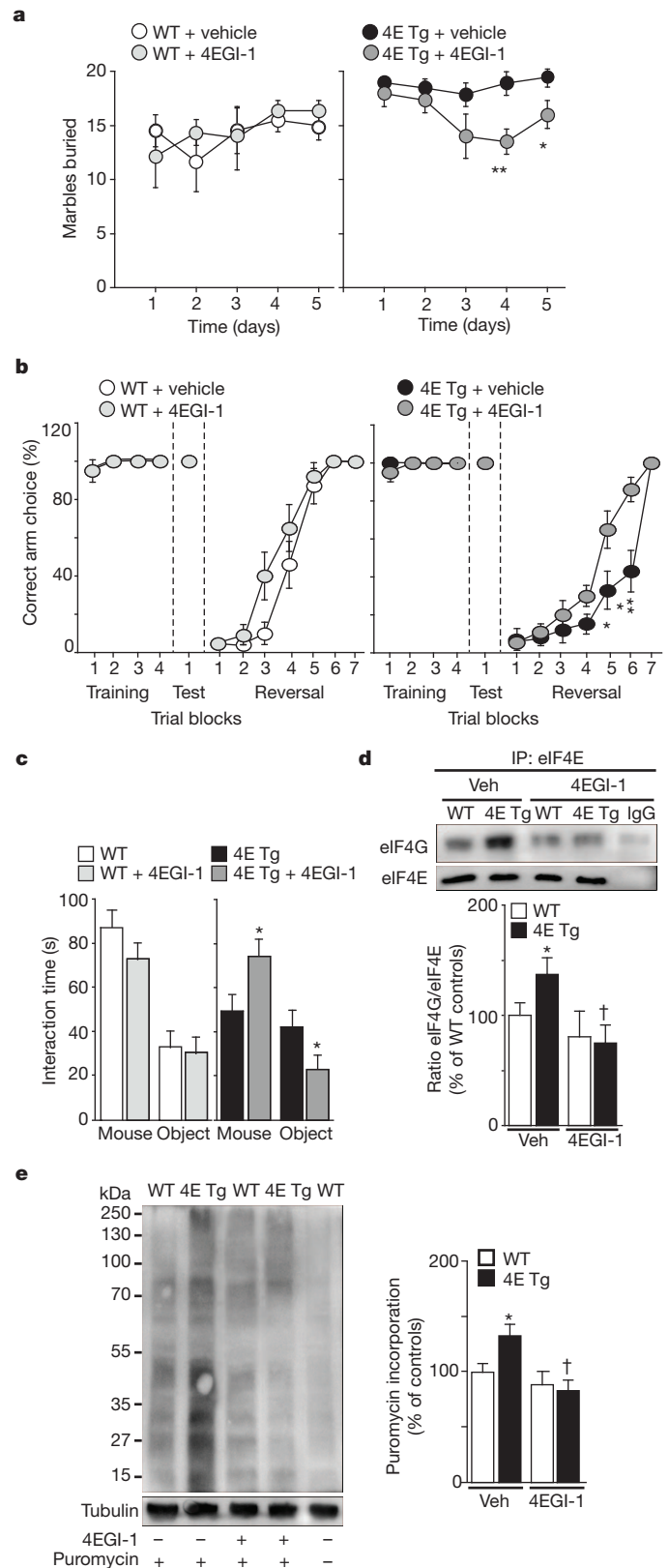
Figure 4 | The cap-dependent translation inhibitor 4EGI-1 reverses ASD-like behaviours shown by eIF4E-transgenic mice. a, Treatment of eIF4E-transgenic mice with 4EGI-1 reduces the marble-burying behaviour. $n = 6$ mice per genotype and treatment; *$P < 0.05$, **$P < 0.01$, two-way repeated-measures ANOVA (treatment $\times$ genotype, $F_{(1,20)} = 4.21$, $P < 0.05$) followed by Bonferroni–Dunn post-hoc test. **b**, 4EGI-1 improves the cognitive flexibility of eIF4E-transgenic mice in the Y-maze test. $n = 6$–7 mice per genotype and treatment; *$P < 0.05$, **$P < 0.01$, two-way repeated-measures ANOVA (treatment $\times$ genotype, $F_{(1,21)} = 4.61$, $P < 0.05$) followed by Bonferroni–Dunn post-hoc test. **c**, 4EGI-1 improves social behaviour of eIF4E-transgenic mice in the three-chamber arena test. $n = 6$ mice per genotype and treatment; *$P < 0.05$, two-way repeated-measures ANOVA (treatment $\times$ genotype, $F_{(1,20)} = 6.26$, $P < 0.05$) followed by Bonferroni–Dunn post-hoc test. **d**, 4EGI-1 decreases the enhanced eIF4E–eIF4G interactions in eIF4E-transgenic mice. Immunoprecipitation of eIF4E in the striatum. $n = 4$ mice per genotype; *$P < 0.05$ (versus vehicle-treated wild type), †$P < 0.05$ (versus 4EGI-1-treated wild type), two-way ANOVA, followed by Bonferroni–Dunn post-hoc test. **e**, 4EGI-1 normalizes the exaggerated cap-dependent translation in eIF4E-transgenic mice as measured with SUnSET. The last wild-type sample represents a control without puromycin. *$P < 0.05$, †$P < 0.05$, two-way ANOVA followed by Bonferroni–Dunn post-hoc test. All data are shown as mean and s.e.m.

Finally, we asked whether exaggerated cap-dependent translation was responsible for the synaptic alterations and ASD-like behaviours shown by the eIF4E-transgenic mice. We took advantage of 4EGI-1, an inhibitor of eIF4E–eIF4G interactions[8,11], to block the synaptic and behavioural consequences of increased eIF4E expression. Bath application of 4EGI-1 normalized the enhanced striatal LTD observed in the eIF4E-transgenic mice (Fig. 3g, h), suggesting that exaggerated striatal LTD (Fig. 3h) is a direct consequence of increased binding of eIF4E to eIF4G (Supplementary Fig. 3i–k).

Next, we used a subthreshold dose of 4EGI-1 (ref. 11) to normalize the behavioural abnormalities in eIF4E-transgenic mice without impairing their wild-type littermates. eIF4E-transgenic mice treated with 4EGI-1 exhibited a decrease in repetitive behaviour during the marble-burying task, which started on day four and persisted throughout day five (Fig. 4a). Moreover, we found that 4EGI-1 maintained the behavioural effects observed in the marble-burying task (Supplementary Fig. 4a, b). We also found that blockade of eIF4E–eIF4G interactions with 4EGI-1 significantly improved the performance of eIF4E-transgenic mice in the reversal phase of the Y-maze test (Fig. 4b). These findings indicate that chronic treatment of eIF4E-transgenic mice with 4EGI-1 reverses their repetitive and perseverative behaviours. We also found that infusions of 4EGI-1 rescued the social behaviour deficits shown by the eIF4E-transgenic mice in the three-chamber arena test, as they exhibited an increased preference for a nonspecific stranger compared to a new object (Fig. 4c).

At the completion of the behavioural studies with 4EGI-1, we performed co-immunoprecipitation experiments, confirming that 4EGI-1 reduced the increased eIF4E–eIF4G interactions exhibited by the eIF4E-transgenic mice (Fig. 4d and Supplementary Fig. 4c–e). Furthermore, puromycin-labelling of newly synthesized proteins was reduced to wild-type levels, indicating that 4EGI-1 was effective in attenuating the increased cap-dependent translation in the eIF4E-transgenic mice (Fig. 4e and Supplementary Fig. 4f, g). Together, these results indicate that repeated treatment of eIF4E-transgenic mice with 4EGI-1 reverses the increased binding of eIF4E to eIF4G, exaggerated cap-dependent translation, and reversal of ASD-like behaviours.

Here we have demonstrated that increased eIF4E expression and, consequently, dysregulated translational control at the initiation phase of protein synthesis in mice results in the appearance of synaptic dysfunction and aberrant behaviours consistent with ASDs. On the basis of our observations, we propose that exaggerated cap-dependent protein synthesis in the eIF4E-transgenic mice and fragile X syndrome model mice[27,28] results in enhanced translation of a specific subset of messenger RNAs. Thus, the identity of both these mRNAs and the *cis*-acting elements in the 5′ untranslated region responsible for eIF4E-dependent protein synthesis and their possible overlap with

fragile X mental retardation protein target mRNAs will be important investigations in future studies.

Our studies with eIF4E-transgenic mice indicate that ASD-like behaviours can be induced by exaggerated cap-dependent translation in the brain. Moreover, we demonstrated that aberrant repetitive, perseverative and social behaviours shown by eIF4E-transgenic mice are reversed by reducing eIF4E–eIF4G interactions, thereby restoring translational homeostasis. Thus, our findings establish a causal link between exaggerated cap-dependent translation and behaviours associated with autism. Finally, our findings indicate that behavioural defects caused by exaggerated cap-dependent translation, which also occurs in fragile X syndrome[29,30], a disorder with a high incidence of autism, are not irrevocable and can be corrected well into adulthood.

METHODS SUMMARY

All procedures involving animals were approved by the New York University Animal Care and Use Committee and followed the National Institutes of Health Guidelines for the use of animals in research. For a detailed description of all the techniques used in this study, please see the Methods. All the experiments were performed with the examiners blinded to genotype.

Full Methods and any associated references are available in the online version of the paper.

1. Levitt, P. & Campbell, D. B. The genetic and neurobiologic compass points toward common signaling dysfunctions in autism spectrum disorders. *J. Clin. Invest.* **119**, 747–754 (2009).
2. Rapin, I. & Tuchman, R. F. Autism: definition, neurobiology, screening, diagnosis. *Pediatr. Clin. North Am.* **55**, 1129–1146 (2008).
3. Kelleher, R. J. & Bear, M. F. The autistic neuron: troubled translation? *Cell* **135**, 401–406 (2008).
4. The Autism Genome Project Consortium. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nature Genet.* **39**, 319–328 (2007).
5. Yonan, A. L. et al. A genomewide screen of 345 families for autism-susceptibility loci. *Am. J. Hum. Genet.* **73**, 886–897 (2003).
6. Neves-Pereira, M. et al. Deregulation of *EIF4E*: a novel mechanism for autism. *J. Med. Genet.* **46**, 759–765 (2009).
7. Ruggero, D. et al. The translation factor eIF-4E promotes tumor formation and cooperates with c-Myc in lymphomagenesis. *Nature Med.* **10**, 484–486 (2004).
8. Moerke, N. J. et al. Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. *Cell* **128**, 257–267 (2007).
9. Gingras, A. C. et al. Hierarchical phosphorylation of the translation inhibitor 4E–BP1. *Genes Dev.* **15**, 2852–2864 (2001).
10. Schmidt, E. K., Clavarino, G., Ceppi, M. & Pierre, P. SUnSET, a nonradioactive method to monitor protein synthesis. *Nature Methods* **6**, 275–277 (2009).
11. Hoeffer, C. A. et al. Inhibition of the interactions between eukaryotic initiation factors 4E and 4G impairs long-term associative memory consolidation but not reconsolidation. *Proc. Natl Acad. Sci. USA* **108**, 3383–3388 (2011).
12. Thomas, A. et al. Marble burying reflects a repetitive and perseverative behavior more than novelty-induced anxiety. *Psychopharmacology (Berl.)* **204**, 361–373 (2009).
13. Peça, J. et al. Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature* **472**, 437–442 (2011).
14. Hoeffer, C. A. et al. Removal of FKBP12 enhances mTOR-Raptor interactions, LTP, memory, and perseverative/repetitive behavior. *Neuron* **60**, 832–845 (2008).
15. Ehninger, D. et al. Reversal of learning deficits in a $Tsc2^{+/-}$ mouse model of tuberous sclerosis. *Nature Med.* **14**, 843–848 (2008).
16. Moy, S. S. et al. Sociability and preference for social novelty in five inbred strains: an approach to assess autistic-like behavior in mice. *Genes Brain Behav.* **3**, 287–302 (2004).
17. Zhou, J. et al. Pharmacological inhibition of mTORC1 suppresses anatomical, cellular, and behavioral abnormalities in neural-specific Pten knock-out mice. *J. Neurosci.* **29**, 1773–1783 (2009).
18. Kwon, C.-H. et al. Pten regulates neuronal arborization and social interaction in mice. *Neuron* **50**, 377–388 (2006).
19. Fineberg, N. A. et al. Probing compulsive and impulsive behaviors, from animal models to endophenotypes: a narrative review. *Neuropsychopharmacology* **35**, 591–604 (2010).
20. Yizhar, O. et al. Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* **477**, 171–178 (2011).
21. Qiu, A., Adler, M., Crocetti, D., Miller, M. I. & Mostofsky, S. H. Basal ganglia shapes predict social, communication, and motor dysfunctions in boys with autism spectrum disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **49**, 539–551 (2010).
22. Tabuchi, K. et al. A neuroligin-3 mutation implicated in autism increases inhibitory synaptic transmission in mice. *Science* **318**, 71–76 (2007).
23. Calabresi, P., Maj, R., Pisani, A., Mercuri, N. B. & Bernardi, G. Long-term synaptic depression in the striatum: physiological and pharmacological characterization. *J. Neurosci.* **12**, 4224–4233 (1992).

24. Hou, L. *et al.* Dynamic translational and proteasomal regulation of fragile X mental retardation protein controls mGluR-dependent long-term depression. *Neuron* **51,** 441–454 (2006).
25. Huber, K. M., Kayser, M. S. & Bear, M. F. Role for rapid dendritic protein synthesis in hippocampal mGluR-dependent long-term depression. *Science* **288,** 1254–1256 (2000).
26. Auerbach, B. D., Osterweil, E. K. & Bear, M. F. Mutations causing syndromic autism define an axis of synaptic pathophysiology. *Nature* **480,** 63–68 (2011).
27. Ronesi, J. A. *et al.* Disrupted Homer scaffolds mediate abnormal mGluR5 function in a mouse model of fragile X syndrome. *Nature Neurosci.* **15,** 431–440 (2012).
28. Sharma, A. *et al.* Dysregulation of mTOR signaling in fragile X syndrome. *J. Neurosci.* **30,** 694–702 (2010).
29. Dölen, G. *et al.* Correction of fragile X syndrome in mice. *Neuron* **56,** 955–962 (2007).
30. Qin, M., Kang, J., Burlin, T. V., Jiang, C. & Smith, C. B. Postadolescent changes in regional cerebral protein synthesis: an *in vivo* study in the *Fmr1* null mouse. **25,** 5087–5095 (2005).

Supplementary Information is available in the online version of the paper.

METHODS

Housing. Generation of βT-*Eif4e* transgenic mice (eIF4E-transgenic mice) has been described previously[7].

For all the experiments, we made use of littermates derived from crossing heterozygotes. Mice were backcrossed to the N10 generation in C57BL/6J mice. Overall, eIF4E-transgenic mice were viable, fertile and showed no gross anatomical abnormalities in the age range used for this study. eIF4E-transgenic mice and their wild-type littermates were housed in groups of 3–4 animals per cage and kept on a regular 12 h light/dark cycle (7:00–19:00 light period). Food and water were available ad libitum.

Surgery and drug infusion. Mice were anaesthetized (ketamine (100 mg kg^{-1}) and xylazine (10 mg kg^{-1})) and mounted onto a stereotaxic apparatus. Cannulae (26-gauge) were implanted unilaterally at the following coordinates: −0.22 mm anterioposterior, +1 mm mediolateral, and −2.4 mm dorsoventral[31]. Mice were allowed 1 week to recover after the surgery.

The infusions of the eIF4E–eIF4G inhibitor 4EGI-1 were performed as described previously[11]. In brief, 4EGI-1 dissolved in 100% dimethylsulphoxide (DMSO) was diluted in vehicle (0.5% (2-hydroxypropyl)-β-cyclodextrin and 1% DMSO in artificial cerebrospinal fluid (ACSF)). Vehicle or 4EGI-1 (20 μM) was infused over 1 min (0.5 μl min^{-1}; Harvard Apparatus). On the last day of treatment, mice received infusion of 4EGI-1 alone or puromycin (25 μg in 0.5 μl) before 4EGI-1 infusions. All behaviour and tissue dissection occurred 1 h after 4EGI-1 infusions.

Behaviour. The following behavioural tests were performed on male eIF4E-transgenic mice and their wild-type littermates (2–6 months of age) as described previously: novelty induced locomotor activity[32], open field[33], elevated plus maze[33], rotarod[34], prepulse inhibition[33], marble[14], social behaviour[16], direct social interaction[35,36], Y-maze and the Morris water maze[7,35].

For all experiments, mice were acclimated to the testing room 30 min before behavioural training and all behaviour apparatuses were cleaned between each trial with 30% ethanol. The experimenter was blinded to genotype and drug treatment while performing and scoring all behavioural tasks. All behavioural tests were performed starting with the least aversive task first (locomotor activity) and ending with the most aversive (either water-based mazes or extinction of fear memory).

Western blots. Mice were killed by decapitation 1 h after the infusion with either 4EGI-1 alone or 4EGI-1 plus puromycin. The striatum and prefrontal cortex were rapidly dissected, placed on an ice-cold surface, and sonicated in 1% SDS and boiled for 10 min. Aliquots (2 μl) of the homogenate were used for protein determination with a BCA (bicinchoninic acid) assay kit (Pierce, Thermo Scientific). Equal amounts of protein (20 μg) for each sample were loaded onto 10% polyacrylamide gels. Proteins were separated by SDS–PAGE and transferred overnight to polyvinylidene difluoride membranes (Immobilon-Psq, Millipore Corporation). The membranes were immunoblotted with antibodies against eIF4E (1:1,000), eIF4G (1:1,000), eIF4B (1:1,000), eIF4A (1:1,000) and 4E-BP (1:1,000) (Cell Signaling Technology). Antibodies against β-actin and tubulin (1:5,000, Cell Signaling Technology) were used to estimate the total amount of proteins. Detection was based on a horseradish peroxidase (HRP)-conjugated secondary antibody (Promega) and chemiluminescence reagent (ECL or ECL plus; GEHealthcare), and visualized using a Kodak 4000MM imager to obtain pixel density values for the band of interest (Carestream). All images were obtained using maximum sensitivity settings with no binning (0–65 K signal range). No images analysed presented saturating signals for the bands of interest (>65 K greyscale value). The amount of each protein was normalized for the amount of the corresponding β-actin or tubulin detected in the sample.

Immunoprecipitation. Tissue was homogenized in ice-cold lysis immunoprecipitation buffer containing (in mM): 40 HEPES, pH 7.5, 150 NaCl, 10 pyrophosphate, 10 glycerophosphate, 1 EDTA and 0.1% CHAPS, protease inhibitor II, phosphatase inhibitor mixture I, II (Sigma-Aldrich). Cleared homogenate (500 μg) was incubated with either anti-eIF4G (2.5 μg) or anti-eIF4E (2.5 μg) (Bethyl Laboratories) and gently shaken overnight at 4 °C. The antibody–lysate mix was incubated with 75 μl IgG bound to agarose beads (Thermo Scientific). The bead–sample slurry was incubated while rocking at 4 °C overnight. Supernatant was removed and saved, and immunoprecipitates were washed three times in lysis buffer, and once in wash buffer (50 mM HEPES, pH 7.5, 40 mM NaCl, 2 mM EDTA). SDS–PAGE buffer was added to the washed immunoprecipitates, which then were resolved on 4 to 12% gradient gels. Efficiency of the immunoprecipitation was determined by examining the supernatant and wash fractions obtained from the procedure on images obtained from Kodak 4000MM imager (see western

blots section). Band density values for coimmunoprecipitated eFI4E, eIF4G and 4E-BP were normalized to immunoprecipitated eIF4G or eIF4E.

SUnSET. A protein synthesis assay was performed as previously described using the SUnSET method[11]. Puromycin-treated samples were identified on blots using the mouse monoclonal antibody 12D10 (1:5,000 from a 5 mg ml^{-2} stock). Because only a small fraction of the brain proteins were labelled, signal from blots was identified with ECL-Advance (GEHealthcare).

Electrophysiology. Hippocampal (400 μm), prefrontal and striatal slices (300 μm) for electrophysiology were prepared as described previously[24].

Solution to maintain slices. Cutting solution (in mM): 110 sucrose, 60 NaCl, 3 KCl, 1.25 NaH$_2$PO$_4$, 28 NaHCO$_3$, 0.5 CaCl$_2$, 7 MgCl$_2$, 5 glucose and 0.6 ascorbate. ACSF (in mM): 125 NaCl, 2.5 KCl, 1.25 NaH$_2$PO$_4$, 25 NaHCO$_3$, 25 D-glucose, 2 CaCl$_2$ and 1 MgCl$_2$. Slices were incubated at room temperature and then were placed in the recording chamber for additional recovery time of 60 min at 33 °C.

Extracellular recordings. Extracellular fEPSPs were recorded as described previously[23,24]. In all the experiments, baseline synaptic transmission was monitored for at least 20 min before LTD induction. Three trains of high-frequency stimulation (3 s duration, 100 Hz frequency at 20 s intervals) were used to induce LTD in striatal slices[23], and 10 min of incubation with DHPG (50 μM) was used to induce mGluR-dependent LTD in hippocampal slices[24]. The slope of fEPSPs was expressed as a percentage of the baseline average before LTD induction.

Intracellular recordings. Medial prefrontal pyramidal cells were illuminated and visualized using a ×60 water-immersion objective mounted on a fixed-stage microscope (BX61-WI, Olympus), and the image was displayed on a video monitor using a charge-coupled device camera (Hamamatsu). Recordings were amplified by multiclamp 700B and digitized by Digidata 1440 (Molecular Devices). The recording electrode was pulled from a borosilicate glass pipette (3–5 MΩ) using an electrode puller (P-97, Sutter Instruments), filled with an internal solution according to the specific experimental requirement, and patched onto the soma. The series resistance of the patch pipette was compensated ∼70% and re-adjusted before each experiment. A measured liquid junction potential was corrected by adjusting the pipette offset. All voltage–clamp recordings were low-pass filtered at 10 kHz and sampled at 50 kHz.

Internal solution for mEPSCs (in mM): 120 caesium-methane-sulphonate, 10 HEPES, 10 EGTA, 4 MgCl$_2$, 0.4 NaGTP, 4 MgATP, 10 phosphocreatine and 5 QX-314 (pH adjusted to 7.3 with CsOH, 290 mOsm). Bicuculline 50 μM and tetrodotoxin 1 μM (Tocris) were added to the external ACSF bath solution.

Internal solution for mIPSCs (in mM): 140 CsCl, 10 EGTA, 10 HEPES, 2 MgCl$_2$, 2.0 Mg-ATP, 4 Na$_2$-ATP, 0.4 Na$_2$-GTP and 5 QX-314 (pH adjusted to 7.3 with CsOH, 290 mOsm), thus yielding a chloride reversal potential of around 2 mV for the chloride currents. Tetrodotoxin (1 μM), 6,7-dinitroquinoxaline-2,3-dione (DNQX) (40 μM) and D-2-amino-5-phosphonopentanoate (AP5) (50 μM) were added to the ACSF bath solution.

In these conditions, mEPSCs and mIPSCs were recorded in voltage clamp at −70 mV and measured for 120 s and 60 s, respectively.

Dendritic spine morphology. Dendritic spine density experiments were performed as previously described[37,38]. In brief, two-photon imaging was accomplished with a custom microscope and high-resolution stacks ($x = 0.13$ μm, $y = 0.13$ μm, $z = 0.2$ μm per voxel) of dendritic segments throughout the entire cell were taken for morphological analysis in NeuronStudio. Spine-head volume was calculated using a rayburst algorithm. Images were deconvolved before volume measurements using custom routines written in MATLAB (Mathworks).

31. Franklin, K. B. J. & Paxinos, G. *The Mouse Brain in Stereotaxic Coordinates* 3rd edn (Academic, 2007).
32. Errico, F. *et al.* The GTP-binding protein Rhes modulates dopamine signalling in striatal medium spiny neurons. *Mol. Cell. Neurosci.* **37,** 335–345 (2008).
33. Banko, J. L. *et al.* Behavioral alterations in mice lacking the translation repressor 4E-BP2. *Neurobiol. Learn. Mem.* **87,** 248–256 (2007).
34. Borgkvist, A. *et al.* Altered dopaminergic innervation and amphetamine response in adult *Otx2* conditional mutant mice. *Mol. Cell. Neurosci.* **31,** 293–302 (2006).
35. Chévere-Torres, I., Maki, J. M., Santini, E. & Klann, E. Impaired social interactions and motor learning skills in tuberous sclerosis complex model mice expressing a dominant/negative form of tuberin. *Neurobiol. Dis.* **45,** 156–164 (2012).
36. Blundell, J. *et al.* Neuroligin-1 deletion results in impaired spatial memory and increased repetitive behavior. *J. Neurosci.* **30,** 2115–2129 (2010).
37. Dumitriu, D., Rodriguez, A. & Morrison, J. H. High-throughput, detailed, cell-specific neuroanatomy of dendritic spines using microinjection and confocal microscopy. *Nature Protocols* **6,** 1391–1411 (2011).
38. Chalifoux, J. R. & Carter, A. G. GABAB receptors modulate NMDA receptor calcium signals in dendritic spines. *Neuron* **66,** 101–113 (2010).

LETTER

Multidecadal variability in East African hydroclimate controlled by the Indian Ocean

Jessica E. Tierney[1,2], Jason E. Smerdon[2], Kevin J. Anchukaitis[1,2] & Richard Seager[2]

The recent decades-long decline in East African rainfall[1] suggests that multidecadal variability is an important component of the climate of this vulnerable region. Prior work based on analysing the instrumental record implicates both Indian[2] and Pacific[1] ocean sea surface temperatures (SSTs) as possible drivers of East African multidecadal climate variability, but the short length of the instrumental record precludes a full elucidation of the underlying physical mechanisms. Here we show that on timescales beyond the decadal, the Indian Ocean drives East African rainfall variability by altering the local Walker circulation, whereas the influence of the Pacific Ocean is minimal. Our results, based on proxy indicators of relative moisture balance for the past millennium paired with long control simulations from coupled climate models, reveal that moist conditions in coastal East Africa are associated with cool SSTs (and related descending circulation) in the eastern Indian Ocean and ascending circulation over East Africa. The most prominent event identified in the proxy record—a coastal pluvial from 1680 to 1765—occurred when Indo-Pacific warm pool SSTs reached their minimum values of the past millennium. Taken together, the proxy and model evidence suggests that Indian Ocean SSTs are the primary influence on East African rainfall over multidecadal and perhaps longer timescales.

The 2010–2011 drought in the Horn of Africa, by some measures the worst drought in 60 years[3], is a reminder that rainfall in this politically and socioeconomically vulnerable region can fluctuate dramatically. Prevailing La Niña conditions in the tropical Pacific were partly to blame; East African rainfall is teleconnected to the El Niño/Southern Oscillation[4,5] (ENSO), and the Horn of Africa experiences droughts during La Niña events and pluvials during El Niño events. However, it is debated whether the failure of the 'long rains' (the rainy season of March, April and May) in 2011—which exacerbated the drought—is related to decadal variability in the Indo-Pacific region[1] or anthropogenic forcing[1,2,6]. It is critical for us to understand the character and mechanisms that drive decadal to centennial shifts in East African rainfall if we are to evaluate future regional projections of drought frequency and food security, but the short length of the instrumental record fundamentally limits our ability to understand variability on these timescales using observational data alone.

Palaeoclimate records in East Africa from the past millennium have the potential to extend the instrumental record and reveal mechanisms driving low-frequency climate variability. At present, annually resolved, absolutely dated terrestrial archives such as tree rings are sparse or underdeveloped in the region, but numerous lake basins in East Africa provide sedimentary archives with sufficient accumulation rates to resolve past-millennium climate[7]. Lake archives nevertheless have a fundamental limitation: they are not absolutely dated and, if constrained by radiocarbon (^{14}C) dating, can have a relatively large (~50–100-yr) temporal uncertainty due to compounded analytical and calibration errors. This uncertainty can make the identification of shared trends between different site archives challenging, especially within the time frame of the past millennium.

Here we synthesize lacustrine hydroclimatic proxy records from East Africa using a Monte Carlo empirical orthogonal function (MCEOF) approach[8] to develop a spatiotemporal view of regional water balance during the past millennium that accounts for time uncertainty. We use proxy data from seven different lake basins (Fig. 1a; see Methods Summary for proxy selection criteria) that include charcoal, run-off indicators, lake level reconstructions and leaf wax hydrogen isotopes. The leading MCEOF of the proxy data explains $32 \pm 6\%$ (2σ) of the variance in the data and describes the primary regional mode of hydroclimatic variance during the past ~700 years (Fig. 1b). For sites that load positively on MCEOF1, the mode describes a pattern of late-medieval drought (1300 to 1400) followed by a gradual transition towards wetter conditions, culminating in peak pluvial conditions from about 1700 to 1750, and a subsequent abrupt transition back towards drier conditions that persist until modern times (1950; Fig. 1b). Sites that load negatively on MCEOF1 show the opposite pattern and are dry during the eighteenth-century pluvial.

Some of the major features of MCEOF1, such as the medieval drought and the eighteenth-century pluvial, have been discussed and identified in previous site-specific studies[9,10] and are also evident in some of the individual proxy data time series (Fig. 1a). Our analysis, however, identifies the spatial patterns of these features as well as their associated uncertainties. The spatial loading patterns of MCEOF1 clearly indicate that the sites closer to the eastern coastal/Horn of Africa region vary in antiphase with the interior Rift Valley sites (Fig. 1c). Although EOF analysis imposes a constraint of orthogonality that can complicate the interpretation of loading patterns in a climatic sense[11], this Horn–Rift dichotomy is consistent with convergence anomalies that occur in East Africa in response to altered Indian Ocean SST gradients[12], and also approximates the spatial pattern of the first EOF of 10-yr low-pass-filtered instrumental rainfall data in the region (Fig. 1 and Supplementary Information). These features suggest that the loading pattern in MCEOF1 represents a real and important aspect of climate variability in East Africa that has dominated hydroclimate during the past millennium on the multidecadal timescale.

To explore the possible climatic mechanisms driving this variability, we analyse the relationship between annual precipitation in easternmost Africa—the coastal area that loads positively on MCEOF1—and SSTs in millennium-long control simulations conducted with fully coupled, atmosphere–ocean general circulation models (AOGCMs). These simulations provide sufficient degrees of freedom to investigate unforced, decadal to centennial climate variability, something that cannot be achieved with instrumental data or model simulations spanning only a few centuries. We use the 1,300- and 3,000-yr control runs from the US National Center for Atmospheric Research (NCAR) CCSM4[13] AOGCM and the Geophysical Fluid Dynamics Laboratory (GFDL) CM2.1[14,15] AOGCM, respectively. Both models correctly simulate the ENSO-driven teleconnection between tropical Indo-Pacific SSTs and East African rainfall apparent in the instrumental record, although the correlation is stronger in the former than in the latter (Supplementary Fig. 3). To isolate low-frequency relationships, we apply a 50-yr low-pass filter to both the SST and the precipitation field, in line with the highest frequency recovered in MCEOF1 (Supplementary Fig. 4), and calculate the field correlations (Fig. 2). We find that although the central and eastern Pacific influence on East African rainfall is strong in the annual fields, it disappears on

[1]Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA. [2]Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York 10964, USA.
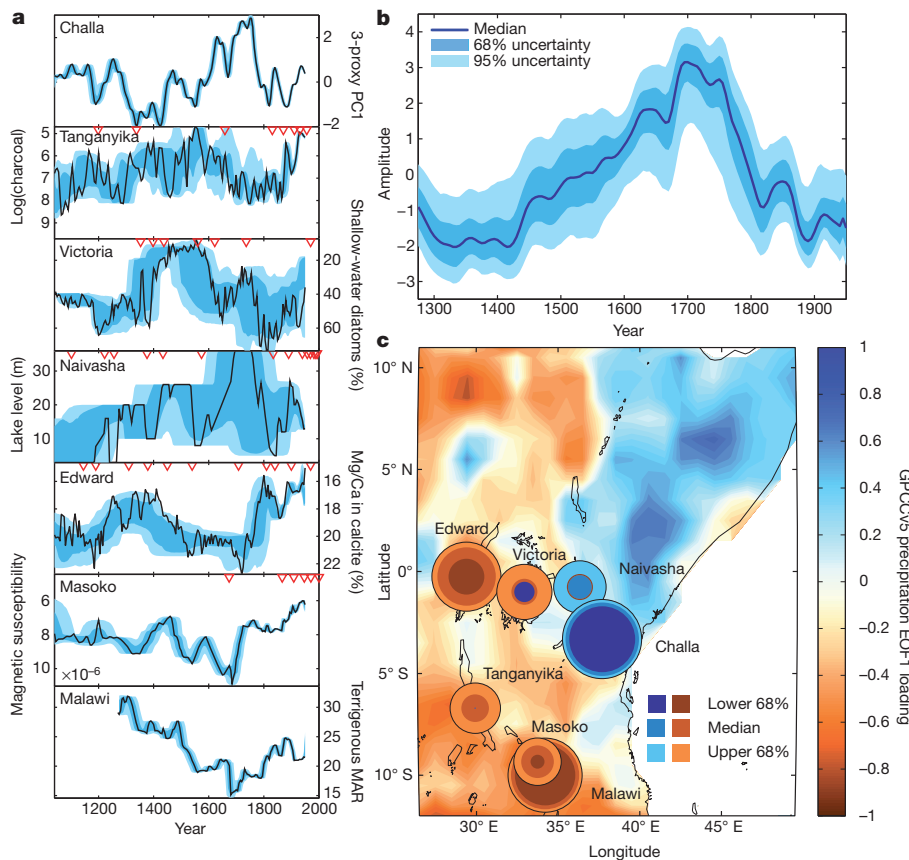
Figure 1 | Results of the MCEOF proxy synthesis. a, The seven proxy data time series used[9,10,20–28] (black lines indicate data on their published age models) and their associated time uncertainties (blue shading: light, 68%; dark, 95%). Red triangles denote chronological constraints (except for the varved Malawi and Challa data). Lake Challa's '3-proxy PC1' is the first principal component of three hydroclimate proxies (Supplementary Information). Vertical axes are oriented such that wet conditions plot upwards. MAR, mass accumulation rate. **b**, MCEOF1, shown with the median (solid blue line) and 68% and 95% two-tailed uncertainty bounds (blue shadings) empirically determined by 10,000 simulations. **c**, Proxy loadings on MCEOF1 (coloured circles), superimposed on the loading pattern (background colours) of the first EOF of the 10-yr low-pass-filtered instrumental precipitation field (GPCCv5[29]). Circle diameter represents the loading value; circle colour represents the lower 68% bound (inner circle), median (middle circle) and upper 68% bound (outer circle).

decadal and longer timescales (Fig. 2 and Supplementary Fig. 5). The only significant correlations that remain in the 50-yr low-pass-filtered field represent the association of wet conditions with warm SSTs in the western Indian Ocean and cool SSTs in the eastern Indian Ocean and western Pacific warm pool (Fig. 2).

Changes in model atmospheric vertical velocity are consistent with the changes in SSTs and provide insight into the mechanisms linking the surface ocean to rainfall variability. Figure 3 shows longitude–height cross-sections of vertical velocity correlated with East African precipitation for the GFDL CM2.1 model. For the unfiltered data (Fig. 3a), wet conditions in East Africa are associated with a reorganization of the Walker circulation throughout the tropics, with anomalous rising motion over the central and eastern equatorial Pacific, descending motion over the far western Pacific and eastern Indian Ocean, and another cell of anomalous rising motion over the western Indian Ocean and East Africa. These correlations are typical of ENSO and are consistent with patterns of low-level wind anomalies driven by the SST gradients and associated atmospheric heating anomalies. In contrast, the low-pass-filtered vertical velocity correlations show that on multidecadal and longer timescales the tropical Pacific influence is considerably weaker (Fig. 3b). Instead, East African precipitation anomalies are controlled by a Walker circulation anomaly that is localized over the Indian Ocean, with wet conditions associated with ascending motion over East Africa and the western Indian Ocean and descending motion over the eastern Indian Ocean.

Collectively, these simulations imply that changes in Indian Ocean SSTs or the Indian Ocean SST gradient—and not the Pacific SST gradient—are the dominant influence on East African rainfall on multidecadal and longer timescales. Such changes in SSTs act to weaken or strengthen the Walker circulation in the Indian Ocean basin, thereby causing multidecadal pluvials or droughts, respectively. The Pacific influence on East African rainfall decays beyond the 10-yr timescale (Supplementary Fig. 5), probably reflecting the waning influence of interannual ENSO. Indeed, analysis of the power spectra of the

simulated SSTs in the Pacific and Indian ocean basins, respectively, suggests that the central Pacific Ocean has relatively little spectral power in the multidecadal band whereas both the eastern and western Indian Ocean basins retain relatively more low-frequency power (Supplementary Fig. 6). These low-frequency oscillations are clearly capable of affecting the Walker circulation in the Indian Ocean basin independently of the Pacific, lending credence to, as well as extending, previous work suggesting that the Indian Ocean directly affects East African hydroclimate even in the interannual band[12,16].

Changing Indian Ocean SST gradients may have dictated the evolution of East African hydroclimate during the past millennium. Independent proxy evidence of Indian Ocean SSTs would provide the most salient means of investigating this, but proxy SST records from the Indian Ocean basin over the past millennium are so far limited to coral archives that span at most the past few centuries. However, a recent sedimentary reconstruction from the Makassar strait provides a continuous SST record for the western Pacific warm pool over the past millennium[17] and, given the known influence of the Indonesian Throughflow on southeastern Indian Ocean SSTs[16], provides an indirect estimate of temperature variability in the easternmost portion of the Indian Ocean basin. This sedimentary reconstruction illustrates a remarkable similarity between the temporal evolution of Makassar SSTs and East African climate, with the coldest SSTs of the past millennium coeval with the coastal pluvial evident in MCEOF1 (Fig. 4). These palaeoclimate data support the mechanistic link between eastern Indian Ocean SSTs and East African rainfall seen in the AOGCM control runs, wherein cool SSTs in the eastern Indian Ocean are associated with wet conditions along the East African coast.

Determination of the role of western Indian Ocean SSTs, which are thought to be critical in terms of promoting local convergence anomalies on the interannual basis[12], awaits long and continuous proxy SST data from the western Indian Ocean. However, it is dynamically consistent with our AOGCM results that cooler conditions in the western Pacific warm pool during the late seventeenth century and
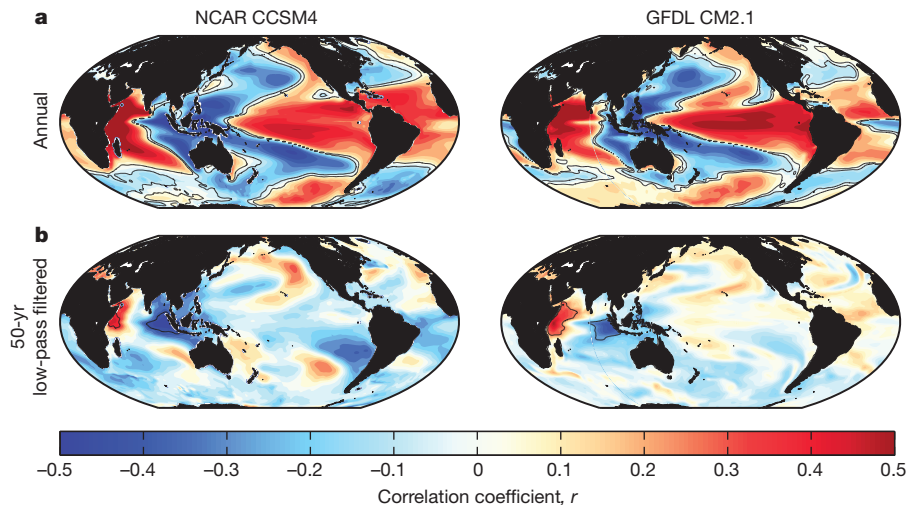
Figure 2 | Field correlations between simulated East African precipitation and SSTs. East African precipitation was averaged over the area spanned by 5° S–7° N and 36°–46° E, and then correlated with SSTs from long control simulations conducted with the NCAR CCSM4 AOGCM and the GFDL CM2.1 AOGCM, respectively. **a**, Annual correlations; **b**, 50-yr low-pass-filtered correlations. Black contour lines demarcate regions in which a null hypothesis of no correlation can be rejected at the 5% level. *r*, Pearson correlation coefficient.

the early eighteenth century helped reduced the east–west gradient in Indian Ocean SSTs, weakened the Walker circulation and promoted pluvial conditions in coastal East Africa and drought in the interior Rift Valley. As the AOGCM simulations demonstrate, such multidecadal modulations of SSTs and East African climate can occur as part of natural, unforced climate variability, leaving open the possibility that the pluvial event in coastal Africa was at least in part a result of internal variability. However, the pluvial event and the cool conditions in the warm pool do occur during one of the coldest intervals of the Northern Hemisphere Little Ice Age[18] and thus may be part of a global climate reorganization in response to radiative forcing.

This work presents an analysis of the spatial and temporal character of East African hydroclimate on decadal to centennial timescales, based on both palaeoclimate data and climate model simulations. In contrast to the dominant impact of ENSO on interannual rainfall variability, variations in the SST gradient across the Indian Ocean—which alter the Walker circulation—seem to be the principal control on hydroclimate in East Africa on multidecadal timescales. Whereas previous interpretations of palaeoclimate data from East Africa typically invoked ENSO or ENSO-like controls on East African precipitation[10], our analysis suggests that palaeoclimate data that preferentially or exclusively record multidecadal or longer climate variability may be better understood in the context of the Indian Ocean, which has its own agency at lower frequencies.

The modelling simulations demonstrate that multidecadal oscillations in Indian Ocean SSTs can arise in response to internal climate
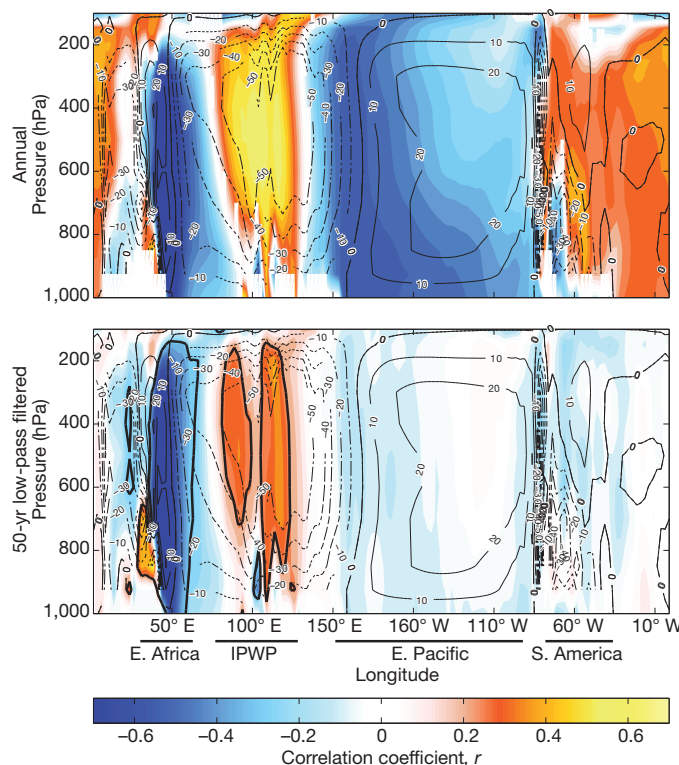

Figure 3 | Field correlations between simulated East African precipitation and vertical velocity. 1Data are from the GFDL CM2.1 simulation. Precipitation was averaged as in Fig. 2 and then correlated with vertical velocity (*ω*; positive values indicate descending motion; averaged over the 5° S–5° N latitude band) across the atmospheric pressure field. **a**, Annual field, showing only values for which a null hypothesis of no correlation can be rejected at the 5% level. **b**, 50-yr low-pass-filtered field. Heavy black contours demarcate regions in which a null hypothesis of no correlation can be rejected at the 5% level. Light contours represent mean-annual simulated *ω* (hectopascals per day), indicating the locations of the ascending and descending branches of the Walker circulation. IPWP, Indo-Pacific warm pool.
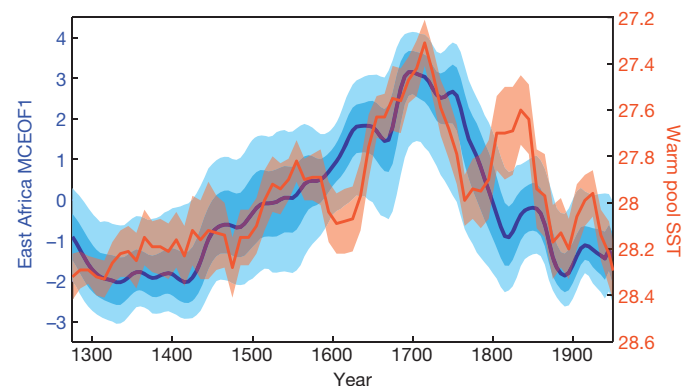

Figure 4 | Comparison between East Africa MCEOF1 (blue) and an SST reconstruction from the Makassar strait in the western Pacific warm pool (orange)[17]. The right-hand axis is flipped such that cooler conditions plot upwards. Shading on MCEOF1 indicates the 68% (dark) and 95% (light) two-tailed uncertainty bounds, as in Fig. 1. Shading on Makassar SSTs indicates the standard error on the binned proxy SST data as described in ref. 17.

variability, but the palaeoclimate data hint at a forced response in that a cooler western Pacific warm pool and a coastal East African pluvial are associated with the Little Ice Age. Efforts to explain the recent decadal drought in the region and project future hydroclimatic change, whether forced or unforced, must acknowledge the existence of potent low-frequency hydroclimate variability related to the Indian Ocean that is not detectable from the instrumental record alone. Present climate models predict that East Africa will get wetter as a consequence of increasing concentrations of greenhouse gases[19], but the region has in fact gotten drier in recent decades[1]. This discrepancy could be explained either if a naturally occurring dry period, possibly of the type identified here, is obscuring a radiatively forced wetting trend or if the model projections are incorrect. More reliable projections of hydroclimate in this vulnerable region will require a better understanding of regional hydroclimate variability on decadal and longer timescales, how the regional climate responds to radiative forcings and how these two factors will combine to determine the future hydroclimate of East Africa.

METHODS SUMMARY

Seven palaeoclimate records of relative water balance from the East African region were included in the MCEOF synthesis, on the basis of the following criteria: the proxy primarily reflects hydroclimate; the proxy was measured at a time interval of 50 years or less; the age model is based on at least seven age–depth tie-points; there is at least one proxy data point representative of modern (post-1950) conditions; and the archive has a well-constrained stratigraphy free of reworking, large turbidites and sedimentary hiatuses. The MCEOF method, which is fully described in ref. 8, was applied to the data with the number of simulations set to 10,000 (ref. 8). Palaeoclimate data and age constraints were treated as described in ref. 8 with the exception of data from Lake Challa, in which case the first principal component of the three available hydroclimate proxy time series was used (Supplementary Information).

The GFDL CM2.1 model's[14] pre-industrial control simulation spans 3,000 years and uses an atmospheric resolution of 2° latitude by 2° longitude by 24 vertical levels and an oceanic resolution of 1° latitude (increasing to 1/3° near the equator) by 1° longitude by 50 vertical levels. The NCAR CCSM4 model's[13] pre-industrial control simulation spans 1,300 years and uses an atmospheric resolution of 1° latitude by 1° longitude by 26 vertical layers and an oceanic resolution of 0.54° latitude (increasing to 0.27° near the equator) by 1.11° longitude by 60 vertical levels. For model analyses, we averaged precipitation over the area spanned by 5° S–7° N and 36°–46° E. Annual precipitation, SST and vertical velocity (for the GFDL simulation) fields were filtered with an eight-point Butterworth filter before calculating the correlation coefficients of the low-pass-filtered fields in Figs 2 and 3. All annual fields were calculated from monthly means. For further details regarding the model analyses, see Supplementary Information.

Data from this paper is available from NOAA's World Data Center for Paleoclimatology (www.ncdc.noaa.gov/paleo).

1. Lyon, B. & DeWitt, D. A recent and abrupt decline in the East African long rains. *Geophys. Res. Lett.* **39,** L02702 (2012).
2. Williams, A. & Funk, C. A westward extension of the warm pool leads to a westward extension of the Walker circulation, drying eastern Africa. *Clim. Dyn.* **37,** 2417–2435 (2011).
3. Famine Early Warning System Network. East Africa: past year one of the driest on record in the eastern Horn. http://www.fews.net/docs/Publications/FEWS NET EA_Historical drought context_061411pdf (2011).
4. Nicholson, S. & Entekhabi, D. The quasi-periodic behavior of rainfall variability in Africa and its relationship to the Southern Oscillation. *Meteorol. Atmos. Phys.* **34,** 311–348 (1986).
5. Ropelewski, C. & Halpert, M. Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Weath. Rev.* **115,** 1606–1626 (1987).
6. Funk, C. *et al.* Warming of the Indian Ocean threatens eastern and southern African food security but could be mitigated by agricultural development. *Proc. Natl Acad. Sci. USA* **105,** 11081–11086 (2008).
7. Verschuren, D. in *Past Climate Variability Through Europe and Africa* (eds Battarbee, R. W., Gasse, F. & Stickley, C. E.) 219–256 (Springer, 2004).
8. Anchukaitis, K. J. & Tierney, J. E. Identifying coherent spatiotemporal modes in time-uncertain proxy paleoclimate records. *Clim. Dyn.* advance online publication, http://dx.doi.org/10.1007/s00382-012-1483-0 (26 August 2012).
9. Verschuren, D., Laird, K. R. & Cumming, B. F. Rainfall and drought in equatorial East Africa during the past 1,100 years. *Nature* **403,** 410–414 (2000).
10. Russell, J. M. & Johnson, T. C. Little Ice Age drought in equatorial Africa: Intertropical Convergence Zone migrations and El Niño-Southern Oscillation variability. *Geology* **35,** 21–24 (2007).
11. Dommenget, D. & Latif, M. A cautionary note on the interpretation of EOFs. *J. Clim.* **15,** 216–225 (2002).
12. Ummenhofer, C. C., Gupta, A. S., England, M. H. & Reason, C. J. C. Contributions of Indian Ocean sea surface temperatures to enhanced East African rainfall. *J. Clim.* **22,** 993–1013 (2009).
13. Gent, P. *et al.* The community climate system model version 4. *J. Clim.* **24,** 4973–4991 (2011).
14. Delworth, T. *et al.* GFDL's CM2 global coupled climate models. Part I: formulation and simulation characteristics. *J. Clim.* **19,** 643–674 (2006).
15. Wittenberg, A. Are historical records sufficient to constrain ENSO simulations? *Geophys. Res. Lett.* **36,** L12702 (2009).
16. Black, E., Slingo, J. & Sperber, K. R. An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian Ocean SST. *Mon. Weath. Rev.* **131,** 74–94 (2003).
17. Oppo, D. W., Rosenthal, Y. & Linsley, B. K. 2,000-year-long temperature and hydrology reconstructions from the Indo-Pacific warm pool. *Nature* **460,** 1113–1116 (2009).
18. Jansen, E. J. *et al.* in *Climate Change 2007: The Physical Science Basis* (eds Solomon, S. *et al.*) 433–498 (Cambridge Univ. Press, 2007).
19. Christensen, J. H. *et al.* in *Climate Change 2007: The Physical Science Basis* (eds Solomon, S. *et al.*) 847–940 (Cambridge Univ. Press, 2007).
20. Verschuren, D. *et al.* Half-precessional dynamics of monsoon rainfall near the East African equator. *Nature* **462,** 637–641 (2009).
21. Tierney, J. E., Russell, J. M., Sinninghe Damsté, J. S., Huang, Y. & Verschuren, D. Late Quaternary behavior of the East African monsoon and the importance of the Congo Air Boundary. *Quat. Sci. Rev.* **30,** 798–807 (2011).
22. Wolff, C. *et al.* Reduced interannual rainfall variability in East Africa during the last ice age. *Science* **333,** 743–747 (2011).
23. Stager, J. C., Ryves, D. B., Cumming, B. F., Meeker, L. D. & Beer, J. Solar variability and the levels of Lake Victoria, East Africa, during the last millenium. *J. Paleolimnol.* **33,** 243–251 (2005).
24. Tierney, J. *et al.* Late-twentieth-century warming in Lake Tanganyika unprecedented since AD 500. *Nature Geosci.* **3,** 422–425 (2010).
25. Garcin, Y. *et al.* Centennial to millennial changes in maar-lake deposition during the last 45,000 years in tropical Southern Africa (Lake Masoko, Tanzania). *Palaeogeogr. Palaeoclimatol.* **239,** 334–354 (2006).
26. Garcin, Y. *et al.* Solar and anthropogenic imprints on Lake Masoko (southern Tanzania) during the last 500 years. *J. Paleolimnol.* **37,** 475–490 (2007).
27. Brown, E. T. & Johnson, T. C. Coherence between tropical East African and South American records of the Little Ice Age. *Geochem. Geophys. Geosyst.* **6,** Q12005 (2005).
28. Johnson, T. & McCave, I. Transport mechanism and paleoclimatic significance of terrigenous silt deposited in varved sediments of an African rift lake. *Limnol. Oceanogr.* **53,** 1622–1632 (2008).
29. Rudolf, B., Becker, A., Schneider, U., Meyer-Christoffer, A. & Ziese, M. GPCC status report, December 2011. http://www.dwd.de/bvbw/generator/DWDWWW/Content/Oeffentlichkeit/KU/KU4/KU42/en/Reports_Publications/GPCC_status_report_2010,templateId=raw,property=publicationFile.pdf/GPCC_status_report_2010.pdf (2011).

Author Contributions J.E.T. designed the study and interpreted the palaeoclimate results. J.E.S. provided the climate modelling results. K.J.A. and J.E.T. designed and implemented the MCEOF method. J.E.S. and R.S. interpreted the dynamical implications of the climate model output. All authors collaborated on the synthesis of the proxy and model data and the writing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.E.T. (tierney@whoi.edu).

LETTER

PKM–ζ is not required for hippocampal synaptic plasticity, learning and memory

Lenora J. Volk[1]*, Julia L. Bachman[1]*, Richard Johnson[1], Yilin Yu[1] & Richard L. Huganir[1]

Long-term potentiation (LTP), a well-characterized form of synaptic plasticity, has long been postulated as a cellular correlate of learning and memory. Although LTP can persist for long periods of time[1], the mechanisms underlying LTP maintenance, in the midst of ongoing protein turnover and synaptic activity, remain elusive. Sustained activation of the brain-specific protein kinase C (PKC) isoform protein kinase M–ζ (PKM–ζ) has been reported to be necessary for both LTP maintenance and long-term memory[2]. Inhibiting PKM–ζ activity using a synthetic zeta inhibitory peptide (ZIP) based on the PKC-ζ pseudosubstrate sequence reverses established LTP _in vitro_ and _in vivo_[3,4]. More notably, infusion of ZIP eliminates memories for a growing list of experience-dependent behaviours, including active place avoidance[4], conditioned taste aversion[5], fear conditioning and spatial learning[6]. However, most of the evidence supporting a role for PKM–ζ in LTP and memory relies heavily on pharmacological inhibition of PKM–ζ by ZIP. To further investigate the involvement of PKM–ζ in the maintenance of LTP and memory, we generated transgenic mice lacking PKC-ζ and PKM–ζ. We find that both conventional and conditional PKC-ζ/PKM–ζ knockout mice show normal synaptic transmission and LTP at Schaffer collateral–CA1 synapses, and have no deficits in several hippocampal-dependent learning and memory tasks. Notably, ZIP still reverses LTP in PKC-ζ/PKM–ζ knockout mice, indicating that the effects of ZIP are independent of PKM–ζ.

PKM–ζ is transcribed from an alternate internal promoter within the _Prkcz_ gene encoding only the catalytic domain, rendering it free from pseudosubstrate auto-inhibition and PKC-ζ regulatory signal transduction pathways[7,8]. To generate mice lacking PKM–ζ we targeted exon 11 of the PKC-ζ catalytic domain for excision by Cre recombinase (Supplementary Fig. 1a). Crossing floxed PKC-ζ alleles (PKC-ζ[fl/fl]) to cytomegalovirus (CMV)-Cre transgenic mice resulted in deletion of both PKC-ζ and PKM–ζ, as verified by southern and western blot analysis (Fig. 1b, c). PKC-ζ/PKM–ζ KO mice were both viable and fertile, and there were no anatomical abnormalities in the brains of PKC-ζ/PKM–ζ KO mice (Fig. 1a). Protein expression levels of a number of synaptic proteins as well as different PKC isozymes, including the other atypical isoform PKC-λ/ι, were unchanged in PKC-ζ/PKM–ζ KO mice (Fig. 1c and Supplementary Fig. 1b–d). Analysis of basal synaptic transmission revealed no difference in input/output relationships in adult PKC-ζ/PKM–ζ KO mice compared to wild-type littermates (Supplementary Fig. 2a, b). Paired-pulse facilitation (PPF) was also unaffected, indicating normal presynaptic release probability in PKC-ζ/PKM–ζ KO mice (Supplementary Fig. 2c). Surprisingly, we found theta-burst stimulation (TBS) resulted in normal induction and maintenance of LTP in PKC-ζ/PKM–ζ KO mice, with a magnitude and time course similar to wild-type littermates (Fig. 1d). Some evidence suggests the dependence of LTP maintenance on PKM–ζ activity can vary with the induction protocol[9]. Therefore, we also tested high frequency stimulation (HFS) and again found no LTP deficit in PKC-ζ/PKM–ζ KO mice (Fig. 1e). As PKM–ζ is proposed to be necessary particularly for the late, protein-synthesis-dependent phase of

LTP[2], we used the protein-synthesis inhibitor emetine to verify that both stimulation protocols produce forms of LTP that are protein-synthesis-dependent in wild-type mice (Supplementary Fig. 3a, b). Furthermore, TBS-LTP in PKC-ζ/PKM–ζ KO mice is also protein-synthesis-dependent on a similar time scale as observed in wild-type mice (Supplementary Fig. 3c).

To limit possible compensatory effects resulting from PKC-ζ/PKM–ζ deletion during development, we also created conditional PKM–ζ knockout mice by crossing PKC-ζ[fl/fl] to CaMKII-CreER[T2] mice. These mice express Cre recombinase which becomes active only upon addition of tamoxifen driven by the _CamkIIα_ (also known as _Camk2a_) promoter[10]. This inducible system allows us to delete PKM–ζ specifically in forebrain neurons of adult animals at will. Adult animals (8–10 weeks old) were injected with tamoxifen (2 mg per day for 5 days) and analysed 2 to 3 weeks later. We observed an ~80% reduction in PKM–ζ levels in the dorsal hippocampus of PKC-ζ[fl/fl]; CaMKII-CreER[T2] mice (Fig. 2a) compared to controls (PKM–ζ[fl/fl] treated with tamoxifen or PKC-ζ[fl/fl]; CaMKII-CreER[T2] treated with vehicle). Remaining PKM–ζ expression is likely limited to interneurons, as evidence suggests PKM–ζ is not expressed in glia[8] (and data not shown). In acute hippocampal slices from conditional PKM–ζ KO mice we saw no differences in input/output curves or PPF (Supplementary Fig. 2d–f), indicating that basal synaptic transmission is unaltered in these animals. Consistent with our findings in conventional PKC-ζ/PKM–ζ KO mice, we found no deficits in either TBS- or HFS-LTP in conditional PKM–ζ KO mice (Fig. 2b, c). As we saw no significant differences between the conventional and conditional knockout animals, we used conventional PKC-ζ/PKM–ζ KO mice for the remainder of our experiments.

Recent studies have raised concerns regarding the specificity and appropriate use of ZIP[11–14]. We find that when added after LTP induction, ZIP was indeed able to reverse LTP in wild-type mice, confirming previously published results[3,15]. In our hands, however, not only were tetanized responses reduced below baseline (see also ref. 16), but ZIP also decreased synaptic transmission in non-tetanized control slices (Fig. 3b). The effects of ZIP were concentration-dependent, as neither LTP maintenance nor baseline transmission was reduced by 1 μM ZIP (Supplementary Fig. 4a–c) and 2–2.5 μM ZIP was effective in roughly half of our experiments (Supplementary Fig. 4d, e), while 5 μM ZIP reliably reduced both LTP and baseline transmission (Fig. 3). The reduction in field excitatory postsynaptic potential (fEPSP) amplitude was not due to changes in pre-synaptic release, as we did not observe changes in PPF following ZIP application (data not shown). Furthermore, whereas a scrambled version of ZIP (Scr-ZIP) is commonly used as a control peptide, some evidence suggests it may mimic the effects of ZIP _in vivo_[17]. To our surprise, Scr-ZIP was equally effective at reversing LTP in acute hippocampal slices (Supplementary Fig. 4f, g). Supporting these results, both peptides inhibit purified PKM–ζ and PKC-ζ activity _in vitro_, showing overlapping inhibition curves (Supplementary Fig. 4h, i). These effects of ZIP and scrambled ZIP were seen with peptides from several different vendors as well as with

[1]Department of Neuroscience, Howard Hughes Medical Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA.
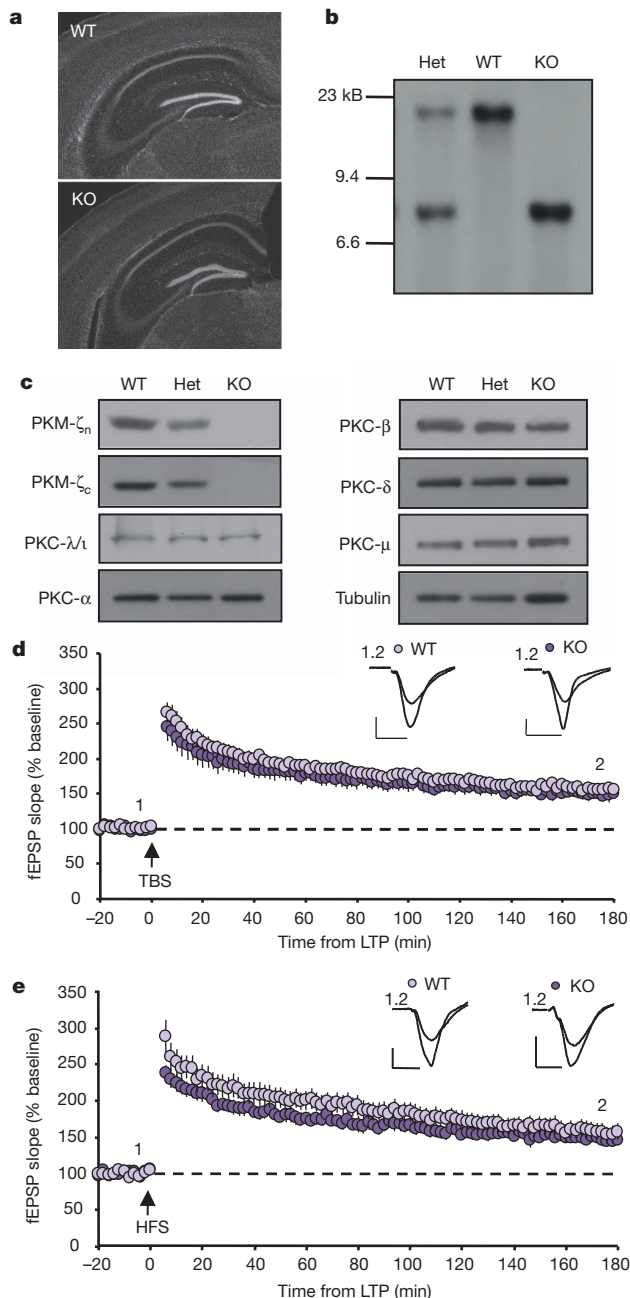*These authors contributed equally to this work.

Figure 1 | Normal LTP, gross brain morphology, and PKC isoform expression in conventional PKC-ζ/PKM-ζ KO mice. a, Hippocampal regions stained with DAPI. **b**, Southern blot analysis of wild-type, heterozygous (Het) and homozygous (KO) PKC-ζ/PKM-ζ mice. **c**, Protein expression of PKC isoforms from whole-brain tissue. **d**, **e**, TBS-LTP (**d**; WT, $n = 11$, $156 \pm 6\%$ at 175–180 min; KO, $n = 9$, $149 \pm 8\%$; $P > 0.5$) and HFS-LTP (**e**; WT, $n = 14$, $155 \pm 11\%$ at 175–180 min; KO, $n = 13$, $148 \pm 6\%$; $P > 0.5$) are intact and maintained for 3 h in mice lacking PKM-ζ. Data represent mean ± s.e.m. Scale bars, 0.5 mV (vertical), 5 ms (horizontal).

Figure 2 | LTP is intact in conditional PKC-ζ/PKM-ζ knockout mice. a, Western blot (left) and quantification (right) of PKM-ζ protein reduction in conditional KO mice normalized to tubulin (WT, $n = 17$, $101 \pm 6.5\%$; KO, $n = 20$, $22 \pm 3.6\%$). **b**, **c**, TBS-LTP (**b**; WT, $n = 16$, $180 \pm 10\%$ at 175–180 min; KO, $n = 16$, $177 \pm 9\%$; $P > 0.8$) and HFS-LTP (**c**; WT, $n = 6$, $154 \pm 11\%$ at 175–180 min; KO, $n = 9$, $148 \pm 10\%$; $P > 0.7$) are intact and maintained for 3 h in mice with PKM-ζ conditionally deleted in adulthood. Data represent mean ± s.e.m. Scale bars, 0.5 mV (vertical), 5 ms (horizontal).

custom synthesized peptides. Importantly, the concentration of ZIP and Scr-ZIP used in our experiments is consistent with that used in previous studies (1–5 μM)[13,15,18].

ZIP and Scr-ZIP use an N-terminal myristoylation group to permeate cell membranes. Several studies in non-neuronal cells have demonstrated that the myristoylation group of ZIP can activate intracellular signalling cascades independent of PKC-ζ activity[19,20]. To control for possible non-specific effects of the myristoylation group, we tested whether myr-PKI, a myristoylated peptide inhibitor of protein kinase A (PKA), would give similar results. Whereas inhibition of PKA
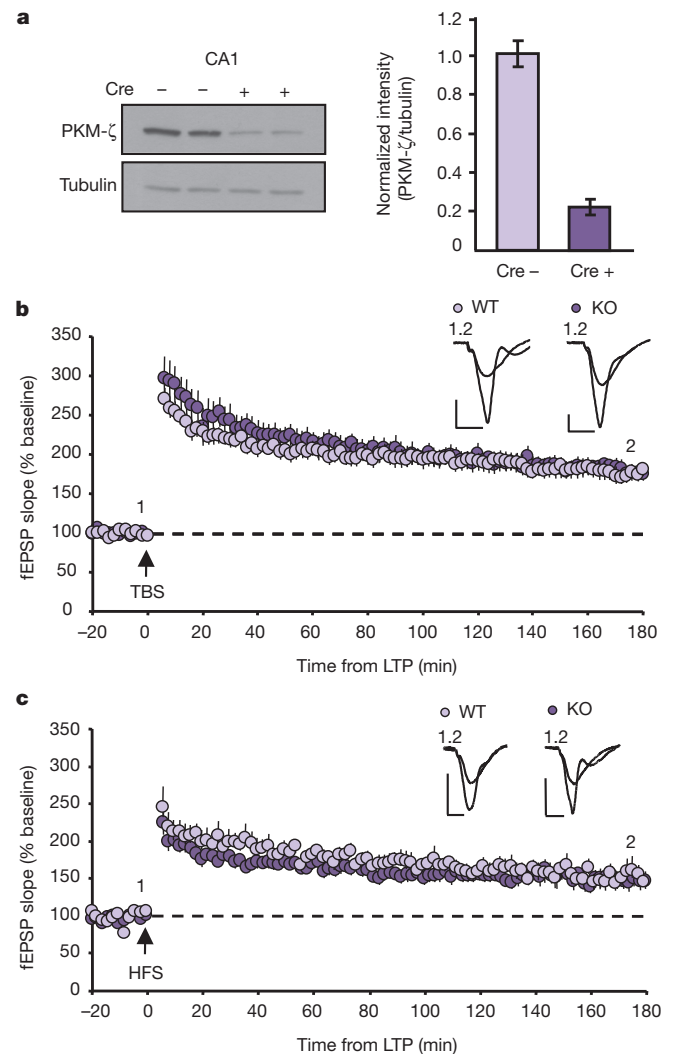
can prevent LTP induction, it has no effect on LTP that is already established[21,22]. Application of myr-PKI 1 h after LTP induction had no effect on LTP maintenance or baseline transmission (Fig. 3d). These data indicate that the reduction of LTP by ZIP is not due to the myristoyl group.

ZIP is reported to be highly specific for PKC-ζ/PKM-ζ[3]. It is therefore expected to have no effect on LTP in PKC-ζ/PKM-ζ KO mice. However, not only did ZIP reverse both TBS- and HFS-LTP in PKC-ζ/PKM-ζ KO mice (Fig. 3a, c), but the efficacy of ZIP in PKC-ζ/PKM-ζ KO mice was indistinguishable from that in wild-type (Fig. 3a–c). This raises the question of whether the other atypical PKC isoform, PKC-λ/ι, which contains the same pseudosubstrate sequence, compensates for the lack of PKM-ζ. To date, there is no evidence that PKC-λ/ι produces a truncated, constitutively active form *in vivo*. To test whether alternate products are generated in the absence of PKM-ζ under LTP conditions, we probed tetanized slices from PKC-ζ/PKM-ζ KO mice with a PKC-λ/ι-specific antibody. We found no evidence for the formation of a truncated product 2 h after LTP induction in PKC-ζ/PKM-ζ KO mice (Supplementary Fig. 5a, b). Expression of full-length PKC-λ/ι was also

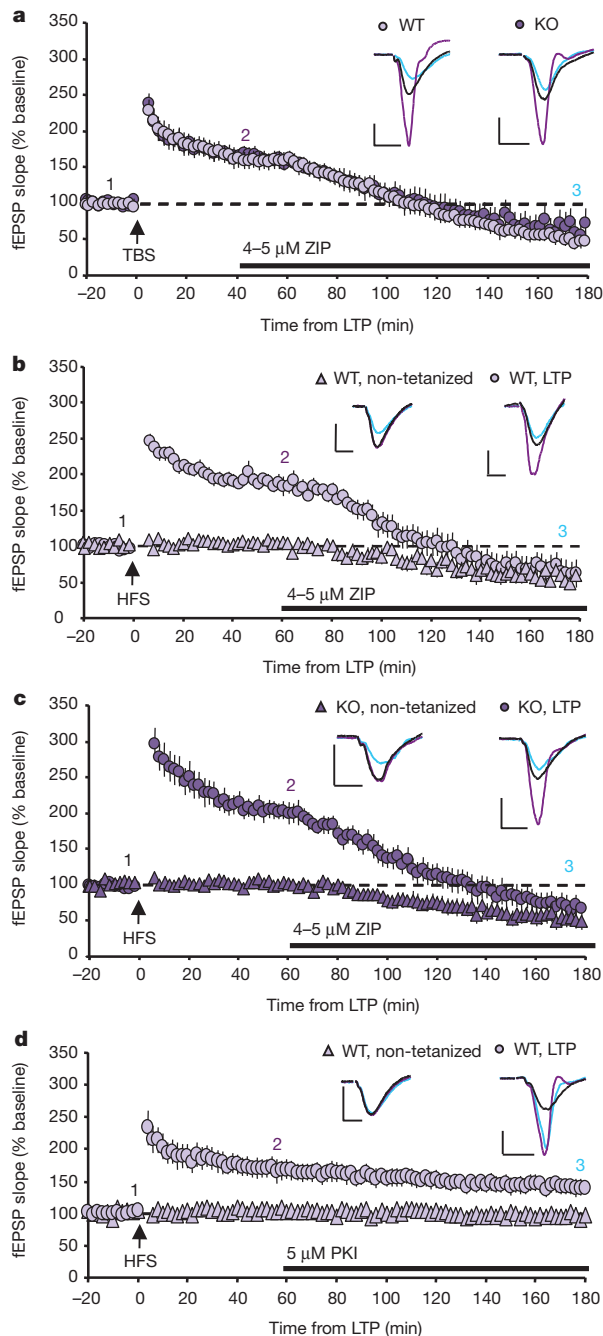**Figure 3 | ZIP is not specific for PKM-ζ. a**, ZIP is equally effective at reversing established TBS-LTP in WT mice and mice lacking PKM-ζ. Before ZIP application (40 min post-LTP): WT, $n = 8$, $166 \pm 12\%$; KO, $n = 6$, $167 \pm 10\%$, $P > 0.9$. 140 min after ZIP application: WT, $48 \pm 7\%$; KO, $60 \pm 18\%$, $P > 0.5$. **b, c**, ZIP decreases both tetanized (HFS) and non-tetanized synaptic responses in WT mice (**b**) and mice lacking PKM-ζ (**c**). Before ZIP application (60 min post-LTP): WT, $n = 8$, tetanized $= 181 \pm 8\%$, non-tetanized $= 102 \pm 4\%$; KO, $n = 8$, tetanized $= 203 \pm 12\%$, non-tetanized $= 100 \pm 4\%$, $P > 0.15$ WT vs KO tetanized or non-tetanized. 120 min after ZIP application: WT, tetanized $= 60 \pm 18\%$, non-tetanized $= 50 \pm 9\%$; KO, tetanized $= 67 \pm 14\%$, non-tetanized $= 52 \pm 11\%$, $P > 0.6$ WT vs KO tetanized or non-tetanized. **d**, Myristoylated PKI peptide does not affect LTP or basal transmission (tetanized $n = 5$, non-tetanized $n = 4$). Data represent mean $\pm$ s.e.m. Scale bars, 0.5 mV (vertical), 5 ms (horizontal).

unchanged. Future work with conditional PKC-λ/ι knockout mice is needed to validate whether PKC-λ/ι has a role in LTP maintenance and could account for the effects of ZIP.

Although the lack of PKM-ζ does not appear to affect canonical forms of LTP in the hippocampus, plasticity mechanisms other than
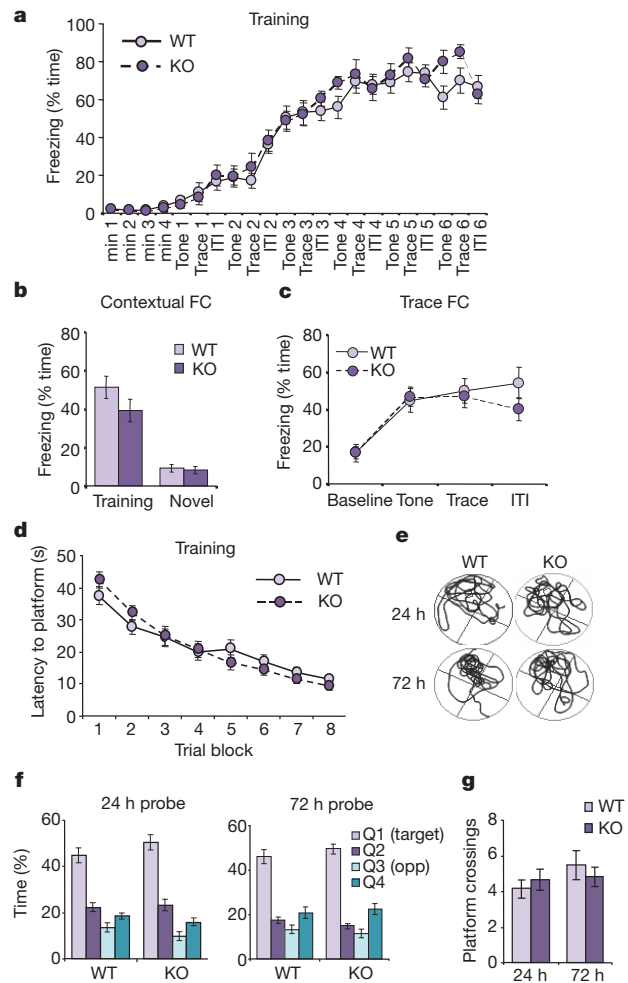

Figure 4 | Hippocampal-dependent learning and memory are intact in PKC-ζ/PKM-ζ KO mice. a, Trace fear conditioning elicits freezing in PKC-ζ/PKM-ζ KO ($n = 13$) and WT ($n = 15$) mice with no significant main effect of genotype ($F_{(1,546)} = 0.50$, $P > 0.48$). ITI, inter-trial interval. **b**, WT and PKC-ζ/PKM-ζ KO mice exhibit similar contextual freezing behaviour (24 h post-training, WT $= 51.5 \pm 5.7\%$, KO $= 39.3 \pm 5.7\%$, $P_{\text{WT vs KO}} > 0.16$; significant elevation in post-training vs pre-training freezing, $P < 0.05$ for WT and KO), but show little freezing in a novel context (48 h post-training, WT $= 9.3 \pm 1.8\%$, KO $= 8.5 \pm 1.9\%$, WT and KO $P > 0.05$ vs pre-training freezing). **c**, Trace fear conditioning is unaffected in PKC-ζ/PKM-ζ KO mice 48 h after training (no significant main effect of genotype, $F_{(1, 78)} = 0.33$, $P > 0.57$). **d**, Mean escape latencies during Morris water maze training (WT $n = 21$, KO $n = 17$, no significant main effect of genotype, $F_{(1,252)} = 0.00$, $P > 0.98$). **e**, Representative swim paths during probe trials for WT and PKC-ζ/PKM-ζ KO mice. **f**, Percentage of time spent in each quadrant during probe trials. Both genotypes showed a significant preference for the target quadrant at 24 (WT $n = 21$, KO $n = 17$) and 72 h (WT $n = 17$, KO $n = 15$); $P < 0.001$ (target quadrant vs Q2, Q3 or Q4). There was no significant main effect of genotype at 24 ($F_{(1,108)} = 0.79$, $P > 0.38$) or 72 h ($F_{(1,90)} = 0.26$, $P > 0.61$). **g**, Number of platform crossings during probe trials (24 h: WT $= 4.14 \pm 0.52$, KO $= 4.64 \pm 0.59$, $P > 0.5$, 72 h: WT $= 5.47 \pm 0.8$, KO $= 4.80 \pm 0.54$, $P > 0.5$). Data represent mean $\pm$ s.e.m.

LTP clearly affect learning and memory[23]. Given the plethora of studies supporting the role of PKM-ζ in memory storage, we tested whether deletion of PKM-ζ could affect two different forms of hippocampal-dependent learning and memory. To assess classical associative conditioning, we chose auditory trace fear conditioning[24]. Freezing behaviour of both wild-type and PKC-ζ/PKM-ζ KO mice increased quickly and markedly during the training session (Fig. 4a). Placing the animals back into the training context 24 h later elicited context-evoked freezing behaviour in PKC-ζ/PKM-ζ KO mice that was not significantly different from wild-type littermates (Fig. 4b). On the

following day, mice were placed in a novel context and presented with only the tone. Whereas neither wild-type or PKC-ζ/PKM-ζ KO mice exhibited significant freezing behaviour in the novel context, tone presentation resulted in similar levels of elevated freezing in both genotypes (Fig. 4b, c). These data show that PKC-ζ/PKM-ζ KO mice exhibit normal contextual and trace components of aversive conditioning.

To test spatial reference memory in PKC-ζ/PKM-ζ KO mice, we evaluated their performance in the Morris water maze, a task where rodents use spatial cues to locate a hidden platform in a pool of opaque water. Latency to find the platform gradually decreased over the course of eight training blocks for both genotypes, indicating PKC-ζ/PKM-ζ KO mice learned the task as well as wild-type mice (Fig. 4d). Memory retention tests were conducted 24 and 72 h after training. Both wild-type and PKC-ζ/PKM-ζ KO mice spent a significantly greater amount of time in the target quadrant (Fig. 4e, f). In addition, the number of times animals crossed through the platform location was not significantly different between the two groups (Fig. 4g). Contrary to results using ZIP infusion in rats[6], our results demonstrate that global as well as specific location memory is unaffected by the deletion of PKM-ζ.

In conclusion, to investigate the role of PKM-ζ in the maintenance of LTP and learning and memory we generated constitutive and conditional PKC-ζ/PKM-ζ knockout mice. Using both prolonged and acute genetic deletion we find that PKM-ζ is not required for the induction or maintenance of LTP in the CA1 region of the hippocampus. The inhibitory peptide ZIP still reversed LTP in the PKC-ζ/PKM-ζ knockout, indicating that ZIP either has non-specific effects or its action is independent of PKM-ζ. Further studies will be necessary to determine ZIP's specificity and site of action. Moreover, deletion of PKC-ζ/PKM-ζ had no effect on two forms of hippocampal-dependent learning and memory. Intracellular perfusion of PKM-ζ has been shown to enhance synaptic transmission by increasing the number of synaptic AMPA (alpha-amino-3-hydroxy-5-methyl-isoxazole-4-propionic acid)-type glutamate receptors[3,25]. PKM-ζ overexpression has also been shown to enhance memory for conditioned taste aversion[26], suggesting that under some conditions or in select brain regions overexpression of PKM-ζ may regulate synaptic plasticity and memory. However, our results demonstrate that normal synaptic plasticity, learning, and memory can occur in the absence of PKM-ζ and call into question an essential, general role for PKM-ζ in the maintenance of synaptic plasticity and memory.

METHODS SUMMARY

Electrophysiology. Extracellular recordings were made from transverse hippocampal slices (2.5–3.5-month-old mice) submerged in standard ACSF. TBS, 4 trains of 10 bursts at 5 Hz; 1 burst = 4 stimuli at 100 Hz, 10 s inter-train interval. HFS, 4 trains of 100 Hz (1 s duration), 20 s inter-train interval. Stimuli during HFS were set at 75% maximum response. All drugs were bath-applied at the indicated time.

Behaviour. 3–4 month old mice were used for behaviour experiments. Trace fear conditioning was conducted as previously described[27] with the exception that the tone test was conducted on day 4 (48 h after training). A detailed schematic is depicted in Supplementary Fig. 6. Morris water maze training consisted of a total of 40 trials over 5 days, with 4, 60 s trials per session and 2 sessions per day separated by approximately 2 h. On the fifth day the first trial consisted of a 60-s probe trial in which the platform was removed (24-h probe trial). The remaining 8 training trials were conducted after the 24-h probe trial. A second probe trial was performed 72 h after the first probe trial. All animals were housed in standard conditions and treated in accordance with the Johns Hopkins University Animal Care and Use Committee guidelines.

Full Methods and any associated references are available in the online version of the paper.

1. Abraham, W. C., Logan, B., Greenwood, J. M. & Dragunow, M. Induction and experience-dependent consolidation of stable long-term potentiation lasting months in the hippocampus. *J. Neurosci.* **22,** 9626–9634 (2002).

2. Sacktor, T. C. PKMζ, LTP maintenance, and the dynamic molecular biology of memory storage. *Prog. Brain Res.* **169,** 27–40 (2008).
3. Ling, D. S. *et al.* Protein kinase Mζ is necessary and sufficient for LTP maintenance. *Nature Neurosci.* **5,** 295–296 (2002).
4. Pastalkova, E. *et al.* Storage of spatial information by the maintenance mechanism of LTP. *Science* **313,** 1141–1144 (2006).
5. Shema, R., Sacktor, T. C. & Dudai, Y. Rapid erasure of long-term memory associations in the cortex by an inhibitor of PKMζ. *Science* **317,** 951–953 (2007).
6. Serrano, P. *et al.* PKMζ maintains spatial, instrumental, and classically conditioned long-term memories. *PLoS Biol.* **6,** e318 (2008).
7. Hernandez, A. I. *et al.* Protein kinase Mζ synthesis from a brain mRNA encoding an independent protein kinase Cζ catalytic domain. Implications for the molecular mechanism of memory. *J. Biol. Chem.* **278,** 40305–40316 (2003).
8. Hirai, T., Niino, Y. S. & Chida, K. PKCζII, a small molecule of protein kinase Cζ, specifically expressed in the mouse brain. *Neurosci. Lett.* **348,** 151–154 (2003).
9. Sajikumar, S. & Korte, M. Metaplasticity governs compartmentalization of synaptic tagging and capture through brain-derived neurotrophic factor (BDNF) and protein kinase Mζ (PKMζ). *Proc. Natl Acad. Sci. USA* **108,** 2551–2556 (2011).
10. Erdmann, G., Schutz, G. & Berger, S. Inducible gene inactivation in neurons of the adult mouse forebrain. *BMC Neurosci.* **8,** 63 (2007).
11. Lisman, J. Memory erasure by very high concentrations of ZIP may not be due to PKM-zeta. *Hippocampus* **22,** 648–649 (2012).
12. Wu-Zhang, A. X., Schramm, C. L., Nabavi, S., Malinow, R. & Newton, A. C. Cellular pharmacology of protein kinase Mζ (PKMζ) contrasts with its *in vitro* profile. *J. Biol. Chem.* **287,** 12879–12885 (2012).
13. Sacktor, T. C. & Fenton, A. A. Appropriate application of ZIP for PKMζ inhibition, LTP reversal, and memory erasure. *Hippocampus* **22,** 645–647 (2012).
14. Yao, Y. *et al.* Matching biochemical and functional efficacies confirm ZIP as a potent competitive inhibitor of PKMζ in neurons. *Neuropharmacol.* **64,** 37–44 (2013).
15. Serrano, P., Yao, Y. & Sacktor, T. C. Persistent phosphorylation by protein kinase Mζ maintains late-phase long-term potentiation. *J. Neurosci.* **25,** 1979–1984 (2005).
16. Sajikumar, S., Navakkode, S., Sacktor, T. C. & Frey, J. U. Synaptic tagging and cross-tagging: the role of protein kinase Mζ in maintaining long-term potentiation but not long-term depression. *J. Neurosci.* **25,** 5750–5756 (2005).
17. Kwapis, J. L., Jarome, T., Lonegran, M. & Helmstetter, F. Protein kinase Mzeta maintains fear memory in the amygdala but not in the hippocampus. *Behav. Neurosci.* **123,** 844–850 (2009).
18. Mei, F., Nagappan, G., Ke, Y., Sacktor, T. C. & Lu, B. BDNF facilitates L–LTP maintenance in the absence of protein synthesis through PKMζ. *PLoS ONE* **6,** e21568 (2011).
19. Krotova, K. *et al.* Peptides modified by myristoylation activate eNOS in endothelial cells through Akt phosphorylation. *Br. J. Pharmacol.* **148,** 732–740 (2006).
20. Lim, S. *et al.* A myristoylated pseudosubstrate peptide of PKC-ζ induces degranulation in HMC-1 cells independently of PKC-ζ activity. *Life Sci.* **82,** 733–740 (2008).
21. Denny, J. B., Polan-Curtain, J., Rodriguez, S., Wayner, M. J. & Armstrong, D. L. Evidence that protein kinase M does not maintain long-term potentiation. *Brain Res.* **534,** 201–208 (1990).
22. Huang, Y. Y. & Kandel, E. R. Recruitment of long-lasting and protein kinase A-dependent long-term potentiation in the CA1 region of hippocampus requires repeated tetanization. *Learn. Mem.* **1,** 74–82 (1994).
23. Kim, S. J. & Linden, D. J. Ubiquitous plasticity and memory storage. *Neuron* **56,** 582–592 (2007).
24. McEchron, M. D., Bouwmeester, H., Tseng, W., Weiss, C. & Disterhoft, J. F. Hippocampectomy disrupts auditory trace fear conditioning and contextual fear conditioning in the rat. *Hippocampus* **8,** 638–646 (1998).
25. Ling, D. S., Benardo, L. S. & Sacktor, T. C. Protein kinase Mζ enhances excitatory synaptic transmission by increasing the number of active postsynaptic AMPA receptors. *Hippocampus* **16,** 443–452 (2006).
26. Shema, R. *et al.* Enhancement of consolidated long-term memory by overexpression of protein kinase Mζ in the neocortex. *Science* **331,** 1207–1210 (2011).
27. Makuch, L. *et al.* Regulation of AMPA receptor function by the human memory-associated gene KIBRA. *Neuron* **71,** 1022–1029 (2011).

Supplementary Information is available in the online version of the paper.

Author Contributions L.J.V., J.L.B. and R.L.H. designed experiments. L.J.V. and J.L.B performed experiments and analysed data. R.J. designed and generated the knockout animals. Y.Y. assisted in mating and genotyping mouse lines. L.J.V., J.L.B. and R.L.H. wrote the manuscript.

METHODS

Animals. Wild-type and PKC-ζ/PKM-ζ knockout (KO) mice were 129/C57BL6 hybrid background or backcrossed with C57BL6 (N10). No differences were observed between backgrounds or sex in either biochemistry or electrophysiology experiments. Behavioural experiments were performed using only the C57BL6 background. All animals were housed in standard conditions and treated in accordance with the Johns Hopkins University Animal Care and Use Committee guidelines.

Generation of mouse lines: the PKC-ζ/PKM-ζ KO mouse was generated by targeting exon 11 for excision by Cre recombinase. A bacterial artificial chromosome (BAC) genomic clone containing *Prkcz* was obtained from RPCI. A fragment of ~12 kb was subcloned into pBlueScript, and the targeting vector was created by inserting an orphan *loxP* site upstream of exon 11 and a neomycin resistance cassette flanked by Flippase Recognition Target sequences (FRT-neo[R]) and a *loxP* site downstream of exon 11 (Supplementary Fig. 1a). Upon verification of homologous recombination by Southern blot analysis, targeted 129 embryonic stem cells were injected into C57BL6 blastocysts at the Transgenic Facility of the Johns Hopkins University School of Medicine. Southern blot analysis using an outer probe confirmed germline transmission in the F[1] generation of chimaeric mice. Conventional PKC-ζ/PKM-ζ KOs were then generated by breeding to a CMV-Cre transgenic mouse line. For conditional PKM-ζ knockout mice, the FRT-neo[R] cassette was first removed by breeding to a transgenic Flippase mouse line[28]. PKC-ζ/PKM-ζ[fl/fl] animals were then crossed with aCaMKII-CreER[T2] [14]. Conditional knockout of *Prkcz* was induced in adult mice by intraperitoneal injection of tamoxifen (Sigma) dissolved in sunflower seed oil. Control mice were either PKC-ζ/PKM-ζ[fl/fl] injected with tamoxifen or PKC-ζ/PKM-ζ[fl/fl]; aCaMKII-CreER[T2] injected with sunflower seed oil.

Histology. Following deep anaesthesia mice were transcardially perfused with 4% paraformaldehyde (w/v) in phosphate-buffered saline (PBS). Brains were removed and post-fixed for an additional 24 h followed by overnight incubations in 10, 20 and 30% sucrose (w/v) in PBS. Brains were frozen on dry ice and sectioned coronally at 50 μm using a standard microtome. Slices containing the hippocampus were rinsed twice 5 min in PBS, blocked and permeabilized for 3 h at room temperature in PBS containing 0.3% Triton-X-100 (PBST) and 10% normal goat serum (NGS; Vector Laboratories) then incubated for 2 h at room temperature in PBST with 2% NGS and DAPI stain (4′,6-diamidino-2-phenylindole, 1:300; Invitrogen). Tissue was then rinsed twice 10 min in PBST followed by twice 10 min in PBS and mounted using PermaFluor (ThermoScientific). Images were acquired using a Zeiss Axiophot microscope with a ×10 objective.

Biochemistry. Western blot analysis: brains were dissected and submerged in 5 ml of ice-cold RIPA buffer (25 mM Tris pH 7.4, 1% Triton-X-100, 0.5% deoxycholate, 0.1% SDS, 100 mM NaCl, 2 mM EDTA, 2 mM EGTA, 50 mM NaF, 10 mM NaPPi, 1 mM Na orthovanadate) supplemented with 1 μM microcystin (Roche) and complete protease inhibitor cocktail (Roche). Tissue was homogenized using a Dounce-homogenizer, rotated for an additional 30 min at 4 °C then spun at 10,000*g* for 30 min at 4 °C. Protein concentration of the supernatant was determined by a bicinchoninic acid assay (Pierce). Samples were denatured by adding 5× SDS-loading sample buffer and incubated at 95 °C for 10 min before loading for SDS–PAGE. Hippocampal slices were homogenized and lysed directly in 100 μl of 2× SDS-loading sample buffer. Proteins were transferred onto PVDF membranes (Millipore), blocked in Tris-buffered saline (TBS) containing 3% non-fat dry milk and incubated with the indicated antibodies diluted in either 3% non-fat dry milk or 5% BSA in TBS with 0.1% Tween-20. Antibodies used are as follows: monoclonal GluA1 (clone 4.9D) and GluA2 (clone 6A) were generated in-house, as were polyclonal anti-C-terminal PKM-ζ (JH6066), anti-N-terminal PKM-ζ (JH6056), and anti-GluN1 (JH4762) after antisera for each was purchased from Covance. Mouse monoclonal anti-PKC-α, PKC-β, PKC-ι, PKC-λ and anti-GRIP1 were from BD Transduction Laboratories. Anti-PKC-δ and anti-PKC-μ were from Cell Signaling. Anti-tubulin and anti-synaptophysin (SVP-38) were from Sigma and anti-PSD95 (K28/43) was from NeuroMab.

In vitro kinase assays: Glutathione-*S*-transferase (GST)-conjugated PKM-ζ was expressed in HEK cells and purified to near-homogeneity using glutathione Sepharose-4B (GE Healthcare). *In vitro* kinase reactions (50 μl final volume) consisted of: 50 mM Tris (pH 7.4), 5 mM MgCl₂, 10 μM peptide ε (Anaspec), 0.5 μg BSA, 25 μM dithiothreitol, GST–PKM-ζ or GST–PKC-ζ (Stressgen) (within linear range for substrate concentration and reaction time), and varying concentrations of ZIP, Scr-ZIP (Sigma, Tocris, Invitrogen, and Johns Hopkins University School of Medicine Biosynthesis and Sequencing Facility), or Myr-PKI (Tocris). Reactions were initiated by the addition of 50 μM [³²P]ATP (1 μCi per assay) and proceeded for 10 min at 30 °C. Reactions were stopped

by pipetting 40 μl onto Whatman P81 paper, and immediately washed in 0.5% o-phosphoric acid in excess and counted using a Beckman scintillation counter. PKM-ζ activity was measured as the difference between counts incorporated in the presence and absence of enzyme. In the absence of substrate, minimal counts were recorded, indicating purified GST–PKM-ζ was maximally auto-phosphorylated.

Electrophysiology. Slice preparation: mice were anaesthetized with the inhalation anaesthetic isofluorane followed by rapid decapitation. 380-μm transverse hippocampal slices were prepared with a vibratome (Leica VT1200S) after dissection of the hippocampus in ice-cold oxygenated (95% O₂/5% CO₂) dissection buffer containing in mM: 2.6 KCl, 1.25 NaH₂PO₄, 26 NaHCO₃, 211 sucrose, 10 glucose, 0.75 CaCl₂, 7 MgCl₂. Slices were recovered submerged in ACSF (in mM: 125 NaCl, 3.25 KCl, 25 NaHCO₃, 1.25 NaH₂PO₄·H₂O, 11 glucose, 2 CaCl₂, 1 MgCl₂) at 30 °C for at least 2 h before recording.

Extracellular recordings: field excitatory postsynaptic potentials (fEPSPs) were evoked at 0.033 Hz as previously described[27]. Theta burst LTP: 4 trains of 10 bursts at 5 Hz, with each burst consisting of 4 stimuli given at 100 Hz, 10 s inter-train interval. High frequency stimulation: 4 trains of 100 stimuli given at 100 Hz with a 20 s inter-train interval. Stimuli during HFS were set at 75% maximum response. Because we observed effects of ZIP on basal transmission, non-tetanized experiments were conducted in separate slices (from the same animal and the same time/recording chamber as tetanized slices) to rule out the possibility of overlapping pathways in the small mouse stratum radiatum. myr-ZIP, myr-Scr-ZIP, myr-PKI and emetine (Sigma) were bath-applied at the indicated time points.

Data analysis and statistics: all plasticity experiments are presented as responses normalized to the average of the 20 min baseline. Every fourth trace (2-min intervals) is shown in graphs due to size limitations. 5 min averages taken at the indicated time were used to calculate the magnitude of plasticity and for statistical tests. A two-tailed, unpaired student's *t*-test was used for determining significance in synaptic plasticity and input–output experiments. All error bars represent standard error of the mean. Sample traces are averages of 6 consecutive traces (3 min), with the stimulation artefact removed for clarity.

Behaviour. Trace fear conditioning: trace fear conditioning was performed using adult (3–4 month old) male mice with the experimenter blind to genotype. Experimental design is illustrated in Supplementary Fig. 6 and was conducted as previously described[27] with the exception that the tone test was conducted on day 4. Data shown in Fig. 4 are from the first testing block. Percentage of time spent freezing was quantified using automated motion detection software (CleverSys). Genotype differences for training and tone-evoked (trace) fear conditioning were analysed with a two-way repeated measures ANOVA (GraphPad Prism). Different time points within each genotype were analysed with repeated measures one-way ANOVA. Contextual fear conditioning between wild type and KO was analysed using two-tailed, unpaired student's *t*-test. Comparisons of contextual freezing across training sessions/contexts within genotype were made using a repeated measures one-way ANOVA. Bonferroni post-hoc tests were used to make pair-wise comparisons where appropriate. *P* was set at 0.05 for all tests.

Morris water maze: the Morris water maze was performed using adult (3–4 month old) male and female mice with the experimenter blind to genotype. Mice were handled for 5 min each for 5 consecutive days before beginning experiments. The arena consisted of a circular pool (diameter of 120 cm) filled with water that was temperature-regulated (24 °C) and made opaque with non-toxic white tempera paint. A square, plexiglass platform (length of 10 cm) was submerged 1 cm below the surface of the water and four local cues were provided to allow spatial map generation. Mice were trained on a total of 40 trials over 5 days, with 4 trials per session and 2 sessions per day separated by approximately 2 h. Prior to the first training trial, mice were given a single habituation trial without the platform to assess any spatial bias. Trials were 60 s and mice that did not find the platform within that time were guided to the platform by the experimenter. Once on top of the platform, mice were left for an additional 10 s before being removed. Start locations (north, south, east and west) were pseudo-randomized so that each start location was used once per session and the sequence of start locations in any session was never used twice. On the fifth day the first trial consisted of a 60 s probe trial in which the platform was removed (24 h probe trial). The remaining 8 training trials were conducted after the 24 h probe trial. A second probe trial was performed 72 h after the first probe trial. Following the second probe trial, visual and sensorimotor skills were assessed by 6 visible platform trials given in 2 sessions of 3 trials each. The top of the platform was positioned 0.5–1 cm above the surface of the water and the sides of the platform were wrapped in black electrical tape to provide a strong visible cue. The position of the visible platform varied from trial to trial. Tracking and analysis of animal movement was done using the ANY-maze tracking system

(SD instruments). Data were analysed by comparing quadrant preferences and escape latencies averaged across animals within groups (wild type or KO) using one-way repeated measures ANOVA and between groups using two-way repeated measures ANOVA with α set at 0.5. Bonferroni post-hoc tests were used for pairwise comparisons. The number of platform crossings was analysed using a two-tailed, unpaired student's t-test.

28. Rodríguez, C. I. *et al.* High-efficiency deleter mice show that FLPe is an alternative to Cre-*loxP*. *Nature Genet.* **25,** 139–140 (2000).

LETTER

Prkcz null mice show normal learning and memory

Anna M. Lee[1], Benjamin R. Kanter[1], Dan Wang[1], Jana P. Lim[1], Mimi E. Zou[1], Chichen Qiu[1], Thomas McMahon[1], Jahan Dadgar[1], Sarah C. Fischbach-Weiss[1] & Robert O. Messing[1]

Protein kinase M-ζ (PKM-ζ) is a constitutively active form of atypical protein kinase C that is exclusively expressed in the brain and implicated in the maintenance of long-term memory[1–9]. Most studies that support a role for PKM-ζ in memory maintenance have used pharmacological PKM-ζ inhibitors such as the myristoylated zeta inhibitory peptide (ZIP) or chelerythrine. Here we use a genetic approach and target exon 9 of the *Prkcz* gene to generate mice that lack both protein kinase C-ζ (PKC-ζ) and PKM-ζ (*Prkcz*[−/−] mice). *Prkcz*[−/−] mice showed normal behaviour in a cage environment and in baseline tests of motor function and sensory perception, but displayed reduced anxiety-like behaviour. Notably, *Prkcz*[−/−] mice did not show deficits in learning or memory in tests of cued fear conditioning, novel object recognition, object location recognition, conditioned place preference for cocaine, or motor learning, when compared with wild-type littermates. ZIP injection into the nucleus accumbens reduced expression of cocaine-conditioned place preference in *Prkcz*[−/−] mice. *In vitro*, ZIP and scrambled ZIP inhibited PKM-ζ, PKC-ι and PKC-ζ with similar inhibition constant (K$_i$) values. Chelerythrine was a weak inhibitor of PKM-ζ (K$_i$ = 76 μM). Our findings show that absence of PKM-ζ does not impair learning and memory in mice, and that ZIP can erase reward memory even when PKM-ζ is not present.

PKM-ζ is a constitutively active atypical kinase that is transcribed from an internal promoter in the *Prkcz* gene[10]. PKM-ζ and PKC-ζ show complementary patterns of expression, with PKM-ζ mainly expressed in the brain and PKC-ζ primarily expressed outside of the nervous system[10]. We used homologous recombination to target exon 9 of the *Prkcz* gene, which encodes the purine-binding site in the catalytic domain of PKC-ζ and PKM-ζ, to generate mice that lack both kinases. We confirmed the effect of *Prkcz* gene deletion on protein expression of PKC-ζ and PKM-ζ by western blot analysis. We detected an immunoreactive band at 52 kDa corresponding to PKM-ζ in brain samples from C57BL/6, wild-type and heterozygous *Prkcz*[+/−] mice, but not *Prkcz*[−/−] mice (Fig. 1a). A 70-kDa protein band, corresponding to PKC-ζ, was present in kidney and lung samples from wild-type mice but not from *Prkcz*[−/−] mice (Fig. 1a). These data confirm that both PKC-ζ and PKM-ζ are absent in *Prkcz*[−/−] mice.

PKC-ι is the third member of the atypical PKC subfamily that includes PKC-ζ and PKM-ζ. Although it is expressed in the brain, nothing is known about its role in regulating behaviour. Although female mice had higher levels of brain PKC-ι immunoreactivity than males, levels were similar in *Prkcz*[−/−] and wild-type mice (two-factor ANOVA, F_{sex} (1,22) = 13.56, $P = 0.001$; $F_{genotype}$ (1,22) = 0.292, $P = 0.59$; $F_{sex \times genotype}$ (1,22) = 0.009, $P = 0.9224$; Fig. 1b, c). Because all PKC isoforms require phosphorylation at the activation loop by phosphoinositide-dependent kinase-1 (PDK1) for catalytic activity[11,12], we investigated if there was increased phosphorylation of PKC-ι at this site (T411) in *Prkcz*[−/−] mice. The ratio of phospho-T411-PKC-ι/total PKC-ι immunoreactivity (Fig. 1d) did not differ by sex or genotype (F_{sex} (1,10) = 0.096, $P = 0.76$; $F_{genotype}$ (1,10) = 0.567, $P = 0.47$; $F_{sex \times genotype}$ (1,10) = 1.01, $P = 0.34$), and there was also no genotype difference when we combined male and female data ($t = 0.744$, $P = 0.47$). These results indicate that loss of PKM-ζ and PKC-ζ does not result in a compensatory increase in the abundance of PKC-ι or in PDK1-mediated phosphorylation of PKC-ι. We were able to detect all other PKC isozymes in brain samples except for PKC-η,
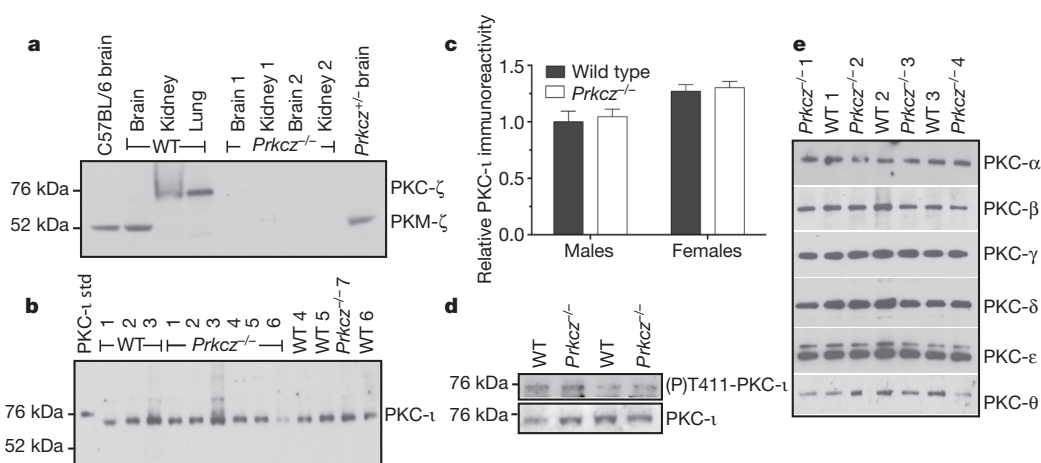

Figure 1 | Absent PKC-ζ and PKM-ζ immunoreactivity in *Prkcz*[−/−] mouse tissues. a, PKC-ζ (∼72 kDa) and PKM-ζ (∼52 kDa) were detected in lung, kidney and brain samples from C57BL/6, wild-type (WT) and heterozygous *Prkcz*[+/−] mice but not from *Prkcz*[−/−] mice. **b**, PKC-ι could be detected in wild-type and *Prkcz*[−/−] samples at ∼72 kDa. His-tagged human PKC-ι (PKC-ι std) was run as a positive control. **c**, Females showed more brain PKC-ι immunoreactivity than males without a difference between genotypes (*n* = 8 wild-type females, *n* = 6 wild-type males, *n* = 5 *Prkcz*[−/−] females, *n* = 7 *Prkcz*[−/−] males). Data are shown as mean + s.e.m. **d**, The phospho-PKC-ι/PKC-ι ratio was similar between *Prkcz*[−/−] (*n* = 7) and wild-type mice brain samples (*n* = 7, *P* = 0.47). **e**, All PKCs, except for PKC-η, were detectable by western blot analysis in wild-type and *Prkcz*[−/−] mouse brain samples, and were of similar abundance in both genotypes.

[1]Ernest Gallo Clinic and Research Center, Department of Neurology, University of California, San Francisco, 5858 Horton Street, Suite 200, Emeryville, California 94608, USA.

and found that their abundance was similar in $Prkcz^{-/-}$ and wild-type mice (Fig. 1e and Supplementary Fig. 1).

$Prkcz^{-/-}$ mice did not show morphological abnormalities or unusual behaviours compared with wild-type mice on a standardized behavioural screen[13]. $Prkcz^{-/-}$ ($n = 11$) and wild-type mice displayed intact visual perception ($n = 22$) on the visual cliff test with both genotypes avoiding the perceived cliff more than 50% of the time and to a similar extent (Mann–Whitney U-test, $U = 116.5$, $P = 0.88$). There was no genotype difference in the tail-flick test for thermal nociception ($n = 13$ per genotype; $t = 0.163$, $P = 0.87$) or in total distance travelled in an open field ($n = 32$ wild type, $n = 30$ $Prkcz^{-/-}$ mice; $t = 0.748$, $P = 0.46$).

We analysed anxiety-like behaviour using the light–dark box and the elevated plus maze, which exploit the conflict between the desire to explore a novel environment and aversion to brightly lit, open spaces. In the light–dark box test, male $Prkcz^{-/-}$ mice spent 51% more time ($F_{genotype \times sex}$ (1,52) = 6.567, $P = 0.01$) and travelled 36% farther in the lit compartment ($F_{genotype \times sex}$ (1,51) = 5.803, $P = 0.02$) than male wild-type mice (Fig. 2a, b). The number of zone crossings was not different between genotype or sex ($F_{genotype}$ (1,52) = 0.560, $P = 0.46$; F_{sex} (1,52) = 1.184, $P = 0.28$; $F_{genotype \times sex}$
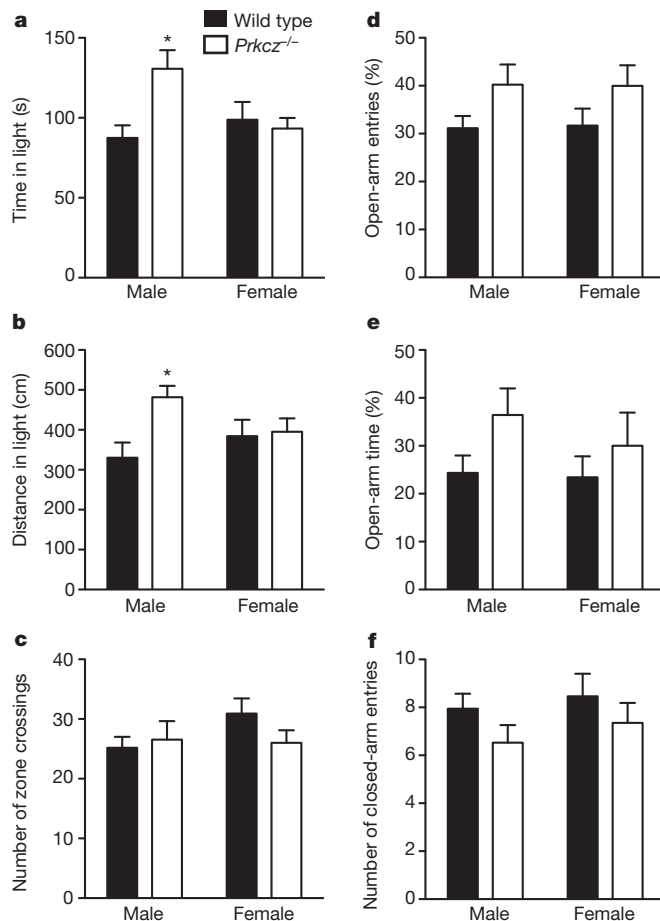

Figure 2 | Reduced anxiety-like behaviour in $Prkcz^{-/-}$ mice. a, b, Male $Prkcz^{-/-}$ mice ($n = 16$) spent more time and travelled farther in the lit compartment compared with male wild-type mice ($n = 13$). There was no genotype difference in female mice ($n = 14$ wild-type mice, $n = 13$ $Prkcz^{-/-}$ mice). *$P < 0.01$ compared with male wild-type mice by Bonferroni post tests. **c,** The total number of zone crossings, a measure of locomotor activity, was similar between $Prkcz^{-/-}$ and wild-type mice of both sexes. **d, e,** $Prkcz^{-/-}$ mice ($n = 31$) made more open-arm entries and showed a trend towards spending more time in the open arms than wild-type mice ($n = 33$). **f,** The number of closed-arm entries was similar in both genotypes. Black bars represent wild-type mice; white bars represent $Prkcz^{-/-}$ mice; data are shown as mean + s.e.m.

(1,52) = 1.725, $P = 0.19$) (Fig. 2c). On the elevated plus maze, $Prkcz^{-/-}$ mice of both sexes ($n = 31$) made more entries into the open arms ($F_{genotype}$ (1,60) = 5.615, $P = 0.02$; F_{sex} (1,60) = 0.002, $P = 0.97$; $F_{genotype \times sex}$ (1,60) = 0.013, $P = 0.91$) and tended to spend more time in the open arms ($F_{genotype}$ (1,60) = 3.302, $P = 0.07$; F_{sex} (1,60) = 0.514, $P = 0.48$; $F_{genotype \times sex}$ (1,60) = 0.285, $P = 0.60$) than wild-type mice ($n = 33$) (Fig. 2d, e). The number of closed-arm entries was not different between genotypes or sexes ($F_{genotype}$ (1,60) = 2.629, $P = 0.11$; F_{sex} (1,60) = 0.752, $P = 0.39$; $F_{genotype \times sex}$ (1,60) = 0.039, $P = 0.85$) (Fig. 2f). These findings indicate that $Prkcz$ modulates anxiety-like behaviour, particularly in male mice.

Because targeting the $Prkcz$ gene reduced anxiety-like behaviour, we also examined fear responses using a cued fear-conditioning procedure, by which animals learn to associate a tone with a foot shock. In this paradigm, subsequent presentations of the tone alone evoke defensive freezing behaviour. Performance in this task is impaired in rats after ZIP administration into the amygdala[7]. We first compared the response to different shock intensities and found no genotype difference ($F_{genotype}$ (1,108) = 0.848, $P = 0.37$; F_{shock} (6,108) = 152.3, $P < 0.0001$; $F_{genotype \times shock}$ (6,108) = 1.298, $P = 0.26$). Automated scoring correlated well with hand scoring for the first tone presentation on the training day ($r^2 = 0.863$, $P < 0.0001$; Supplementary Fig. 2a), the average of five tone presentations on test day 1 ($r^2 = 0.894$; $P < 0.0001$; Supplementary Fig. 2b) and the average of three tone presentations on test day 2 ($r^2 = 0.786$; $P < 0.0003$; Supplementary Fig. 2c). There was a low level of freezing during the first tone presentation on the training day when the mice had not yet been exposed to the shock (Fig. 3a). On test day 1, freezing to the tone was significantly greater, indicating that mice had learned to associate the tone with the foot shock. The mice exhibited a similar level of freezing on test day 2. We found no genotype difference in freezing during any of the sessions, indicating that $Prkcz^{-/-}$ and wild-type mice learned the association equally well ($F_{genotype}$ (1,90) = 0.070, $P = 0.79$; $F_{session}$ (2,90) = 134.6, $P < 0.0001$; $F_{genotype \times session}$ (2,90) = 0.30, $P = 0.745$).

Given that fear memory was unimpaired in $Prkcz^{-/-}$ mice, we investigated other tests of learning and memory. We first used a novel object recognition task to examine hippocampal-dependent learning and memory[14]. Wild-type ($t = 3.69$, $P = 0.002$) and $Prkcz^{-/-}$ mice ($t = 2.98$, $P = 0.01$) showed greater exploration of the novel object compared with chance, and there was no genotype difference in time exploring the novel object ($t = 0.31$, $P = 0.76$) (Fig. 3b). We tested object location memory using a procedure in which performance is impaired in rats administered ZIP bilaterally into the hippocampus[1]. Mice of both genotypes spent more time exploring the new location ($F_{genotype}$ (1,23) = −6.9, $P = 1.00$; $F_{location}$ (1,23) = 9.59, $P = 0.005$; $F_{genotype \times location}$ (1,23) = 0.04, $P = 0.83$) (Fig. 3c). We assessed motor learning by measuring improvement in ability to remain on an accelerating rotarod over successive trials (Fig. 3d). There was no genotype difference in improvement of performance over time ($F_{genotype}$ (1,84) = 0.002, $P = 0.96$; $F_{session}$ (4,84) = 33.29, $P < 0.0001$; $F_{genotype \times session}$ (4,84) = 1.53, $P = 0.20$), indicating that $Prkcz^{-/-}$ and wild-type mice learned this task equally well. We also assessed drug reward memory in male mice by measuring cocaine-conditioned place preference (CPP)[15]. Wild-type ($t = 3.838$, $P = 0.006$) and $Prkcz^{-/-}$ mice ($t = 3.645$, $P = 0.01$) spent significantly more time in the cocaine-paired chamber after conditioning and there was no genotype difference in the cocaine CPP index ($t = 0.153$, $P = 0.88$) (Fig. 3e). Because ZIP injection into the nucleus accumbens can erase cocaine reward memory in rats[3], we tested whether ZIP reduces cocaine reward memory in male and female $Prkcz^{-/-}$ mice using the same cocaine treatment protocol[3]. We found that compared with saline, ZIP impaired cocaine CPP in $Prkcz^{-/-}$ mice ($t = 2.258$, $P = 0.04$) (Fig. 3f).

The finding that ZIP inhibits memory in $Prkcz^{-/-}$ mice suggests that its effect on memory maintenance[1–9] occurs through PKM-ζ-independent mechanisms. Recently, the specificity of ZIP and chelerythrine for inhibiting PKM-ζ has been called into question[16–19]. Part of
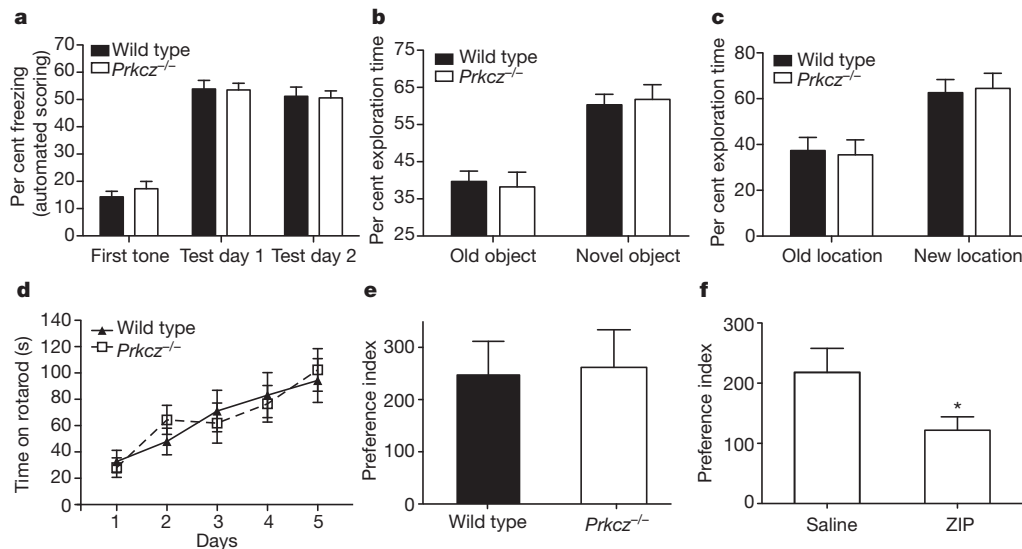
Figure 3 | Intact learning and memory in *Prkcz*⁻/⁻ mice. a, In cued fear conditioning, wild-type (*n* = 20) and *Prkcz*⁻/⁻ mice (*n* = 27) showed similar levels of freezing during all three sessions. **b**, During the novel object task, *Prkcz*⁻/⁻ (*n* = 17) and wild-type mice (*n* = 16) spent more time exploring the new object compared with the old object. There was no genotype difference in exploration of old or novel object. **c**, In the spatial memory task, *Prkcz*⁻/⁻ (*n* = 13) and wild-type mice (*n* = 12) spent more time exploring the new location compared with the old location and there was no genotype difference in exploration of either location. **d**, *Prkcz*⁻/⁻ (*n* = 11) and wild-type mice

(*n* = 12) remained on the accelerating rotarod for similar amounts of time and showed similar improvement in this task over successive trials. **e**, In the CPP test, *Prkcz*⁻/⁻ (*n* = 7) and wild-type mice (*n* = 8) showed similar preference for the cocaine-paired chamber when tested one day after the last conditioning session. **f**, Compared with injection of saline (*n* = 9), injection of ZIP (*n* = 13) into the nucleus accumbens significantly reduced expression of cocaine CPP in *Prkcz*⁻/⁻ mice. All data are shown as mean ± s.e.m. *$P = 0.04$ by two-tailed *t*-test.

this concern arises because the PKC-ζ and -ι pseudosubstrate peptide sequences are identical (SIYRRGARRWRKL), and PKC-ι is widely expressed in the nervous system[20]. To determine the specificity of ZIP, scrambled ZIP and chelerythrine for PKM-ζ, we tested these compounds in an *in vitro* kinase assay using purified PKM-ζ, PKC-ζ and PKC-ι (Table 1 and Supplementary Fig. 3). We found that both ZIP and scrambled ZIP inhibited PKM-ζ in the micromolar range, which is less potent than recently reported[21]. There was only a 7.3-fold difference in K_i values between ZIP and scrambled ZIP, and this modest difference in K_i values suggests that scrambled ZIP is not an ideal control peptide for ZIP inhibition of PKM-ζ. Interestingly, ZIP and scrambled ZIP were equally potent inhibitors of PKC-ι and PKC-ζ compared with PKM-ζ (Table 1). Chelerythrine was a weak inhibitor of PKM-ζ when assayed without dithiothreitol (DTT), and lost all inhibitory activity when 1 mM DTT was included (Table 1 and Supplementary Fig. 3d). These results question the use of ZIP and chelerythrine as specific inhibitors of PKM-ζ.

Our *in vitro* studies indicate that the current pharmacological reagents commonly used to inhibit PKM-ζ are not specific for this kinase. More importantly, our *in vivo* studies with *Prkcz*⁻/⁻ mice indicate that PKM-ζ is not required for long-term memory and that ZIP can impair memory through mechanisms that do not involve PKM-ζ. These findings cast doubt on the importance of PKM-ζ in the maintenance of long-term memory.

METHODS SUMMARY

Generation and testing of *Prkcz*⁻/⁻ mice. A targeting construct containing a 1.1-kilobase (kb) floxed region of exon 9 was used to generate chimaeric mice that were crossed to produce wild-type and *Prkcz*⁻/⁻ littermates. To detect PKC-ζ and PKM-ζ, an anti-PKC-ζ antibody (T. Sacktor) was used. Anti-phospho-PKC (pan) (ζThr 410), which detects PKC-ι phosphorylated at T411, was purchased from Cell Signaling Technology, as were antibodies to detect PKC-α, PKC-δ, PKC-ι and PKC-θ. Anti-PKC-β, PKC-γ and PKC-η antibodies were purchased from BD Transduction Laboratories. Anti-PKC-ε antibody was previously generated (SN134)[22].

Behaviour. The behavioural screen was based on methods described previously[13]. Methods for other behavioural tests are available in the online version of the manuscript. Data were examined for normality using a D'Agostino and Pearson omnibus normality test. Light–dark box and elevated plus maze results were analysed by two-factor analysis of variance (ANOVA) with a Bonferroni post-hoc test. The relationship between machine and hand scoring of fear conditioning was analysed by calculating a Pearson product-moment correlation coefficient. In all tests of learning and memory we did not detect a sex difference; to increase power and the possibility of detecting a genotype difference, we combined results from male and female mice and analysed these data by ANOVA, *t*-, or Mann–Whitney *U*-tests.

Full Methods and any associated references are available in the online version of the paper.

1. Hardt, O., Migues, P. V., Hastings, M., Wong, J. & Nader, K. PKMζ maintains 1-day- and 6-day-old long-term object location but not object identity memory in dorsal hippocampus. *Hippocampus* **20**, 691–695 (2010).
2. He, Y. Y. *et al.* PKMζ maintains drug reward and aversion memory in the basolateral amygdala and extinction memory in the infralimbic cortex. *Neuropsychopharmacology* **36**, 1972–1981 (2011).
3. Li, Y. Q. *et al.* Inhibition of PKMζ in nucleus accumbens core abolishes long-term drug reward memory. *J. Neurosci.* **31**, 5436–5446 (2011).
4. Migues, P. V. *et al.* PKMζ maintains memories by regulating GluR2-dependent AMPA receptor trafficking. *Nature Neurosci.* **13**, 630–634 (2010).
5. Parsons, R. G. & Davis, M. Temporary disruption of fear-potentiated startle following PKMz inhibition in the amygdala. *Nature Neurosci.* **14**, 295–296 (2011).
6. Pastalkova, E. *et al.* Storage of spatial information by the maintenance mechanism of LTP. *Science* **313**, 1141–1144 (2006).

Table 1 | *In vitro* inhibitory activity of compounds against purified atypical PKC isozymes

Compound	PKC-ζ (μM)	PKM-ζ (μM)	PKC-ι (μM)
ZIP	1.70 (1.14–2.54)	2.11 (1.91–2.33)	1.43 (1.21–1.68)
Scrambled ZIP	5.51 (3.12–9.71)	15.4 (14.7–16.2)	4.92 (3.48–6.96)
Chelerythrine	NA	75.97 (68.44–84.25)	NA

Values are K_i, 95% CI. NA, not assessed.

7. Serrano, P. *et al.* PKMζ maintains spatial, instrumental, and classically conditioned long-term memories. *PLoS Biol.* **6,** e318 (2008).
8. Shabashov, D., Shohami, E. & Yaka, R. Inactivation of PKMζ in the NAc shell abolished cocaine-conditioned reward. *J. Mol. Neurosci.* **47,** 546–553 (2011).
9. Shema, R., Sacktor, T. C. & Dudai, Y. Rapid erasure of long-term memory associations in the cortex by an inhibitor of PKM zeta. *Science* **317,** 951–953 (2007).
10. Hernandez, A. I. *et al.* Protein kinase M ζ synthesis from a brain mRNA encoding an independent protein kinase C ζ catalytic domain. Implications for the molecular mechanism of memory. *J. Biol. Chem.* **278,** 40305–40316 (2003).
11. Chou, M. M. *et al.* Regulation of protein kinase C ζ by PI 3-kinase and PDK-1. *Curr. Biol.* **8,** 1069–1078 (1998).
12. Le Good, J. A. *et al.* Protein kinase C isotypes controlled by phosphoinositide 3-kinase through the protein kinase PDK1. *Science* **281,** 2042–2045 (1998).
13. Crawley, J. N. Behavioral phenotyping strategies for mutant mice. *Neuron* **57,** 809–818 (2008).
14. Clarke, J. R., Cammarota, M., Gruart, A., Izquierdo, I. & Delgado-Garcia, J. M. Plastic modifications induced by object recognition memory processing. *Proc. Natl Acad. Sci. USA* **107,** 2652–2657 (2010).
15. Brabant, C., Quertemont, E. & Tirelli, E. Influence of the dose and the number of drug-context pairings on the magnitude and the long-lasting retention of cocaine-induced conditioned place preference in C57BL/6J mice. *Psychopharmacology (Berl.)* **180,** 33–40 (2005).
16. Lisman, J. Memory erasure by very high concentrations of ZIP may not be due to PKM-ζ. *Hippocampus* **22,** 648–649 (2011).
17. Davies, S. P., Reddy, H., Caivano, M. & Cohen, P. Specificity and mechanism of action of some commonly used protein kinase inhibitors. *Biochem. J.* **351,** 95–105 (2000).
18. Wu-Zhang, A. X., Schramm, C. L., Nabavi, S., Malinow, R. & Newton, A. C. Cellular pharmacology of protein kinase Mζ (PKMζ) contrasts with its *in vitro* profile:

implications for PKMζ as a mediator of memory. *J. Biol. Chem.* **287,** 12879–12885 (2012).
19. Nishikawa, K., Toker, A., Johannes, F. J., Songyang, Z. & Cantley, L. C. Determination of the specific substrate sequence motifs of protein kinase C isozymes. *J. Biol. Chem.* **272,** 952–960 (1997).
20. Naik, M. U. *et al.* Distribution of protein kinase Mζ and the complete protein kinase C isoform family in rat brain. *J. Comp. Neurol.* **426,** 243–258 (2000).
21. Yao, Y. *et al.* Matching biochemical and functional efficacies confirm ZIP as a potent competitive inhibitor of PKMζ in neurons. *Neuropharmacology* **64,** 37–44 (2013).
22. Choi, D. S., Wang, D., Dadgar, J., Chang, W. S. & Messing, R. O. Conditional rescue of protein kinase C ε regulates ethanol preference and hypnotic sensitivity in adult mice. *J. Neurosci.* **22,** 9905–9911 (2002).

METHODS

Generation of *Prkcz*^{−/−} mice. A 14.5-kb targeting construct containing exon 9 flanked by *loxP* sites was used to generate ES cells by homologous recombination. Targeted ES cells (W4 line, Taconic) were injected into C57BL/6J blastocysts to generate chimaeric mice that were mated with Flpase C57BL/6J mice (Jackson Laboratories) to remove the neomycin selection cassette in the targeting vector. F_1 generation progeny were crossed with C57BL/6J CMV-Cre mice to delete exon 9. Hybrid C57BL/6JX129S6 wild-type and *Prkcz*^{−/−} littermates were genotyped using the forward primer (GGTATAGTAGGCAGCTATTGCG) located in the long arm of the construct and a reverse primer (TCCTGCCTCAGCCAGAAAAC AAACCACACGG) located outside of the construct. All mice were 8–12-weeks old and housed under a 12-h light:12-h dark cycle, with lights on at 6:00 and off at 18:00. Food and water were freely available. All procedures were conducted in accordance with guidelines of the NIH and the Gallo Center Institutional Animal Care and Use Committee.

Western blotting. Tissue samples were homogenized in RIPA buffer with EGTA, protease and phosphatase inhibitors (G Biosciences). Anti-PKC-ζ (from T. Sacktor) or anti-PKC-ε (SN134[22]) was used at 1:1,000. Anti-PKC-ζ(phospho-Thr 410, catalogue no. 2060), PKC-α, PKC-δ, PKC-ι and PKC-θ antibodies (Cell Signaling Technology) were used at 1:500–1:1,000. Mouse monoclonal anti-PKC-β, PKC-γ and PKC-η antibodies were from BD Transduction Laboratories. All antibodies were incubated in 5% non-fat dry milk, except anti-PKC-ζ(phospho-Thr 410), which was incubated in 5% BSA. HRP-conjugated donkey anti-rabbit or donkey anti-mouse secondary antibodies were used (Jackson Immuno Research Labs). Immunoreactive bands were quantified using ImageJ (http://rsbweb.nih.gov/ij/). Phospho-PKC and PKC samples were normalized to proteins (38–102 kDa) detected on a Coomassie-blue-stained gel run in parallel. Data were expressed relative to the mean immunoreactivity determined in wild-type samples.

Behavioural testing. We examined mice for morphological defects, body weight, and startle to a sudden loud noise[13]. Vision was assessed using a visual cliff assay[13], and the percentage of entries onto the normal perspective surface out of ten trials was calculated. Thermal sensation was tested using a tail-flick apparatus (Columbus Instruments). Locomotor activity was recorded as the distance travelled in an open-field chamber[23,24] after 60 min. Anxiety-like behaviour was measured using a light–dark box (Med Associates Inc.) and an elevated plus maze as in previous work[23,24]. Novel object recognition[14] and object location memory[1] were measured using published methods. Motor learning was assessed using a rotarod (Accuscan Instruments) that accelerated from 0 to 40 r.p.m. in 5 min. Mice were placed on the rotarod at 4 r.p.m., and had to stay on the rotarod for at least 15 s for a trial to be considered successful. The latency to fall in three successful trials was recorded for 5 consecutive days.

The response to foot-shock was determined by administering 0.5-s shocks every 3 min in 0.1-mA increments from 0.1 to 0.7 mA. Responses were scored as: 0, no reaction; 1, flinch; 2, small hop; 3, dash; 4, small jump; 5, large jump, with an extra 0.25 added for any vocalization. Fear conditioning was tested by subjecting naive mice to an 11-min session with five pairings of a 30-s, 85-dB tone that co-terminated with a 1-s, 0.3-mA foot-shock. The chambers (San Diego Instruments) had transparent walls with a metal rod floor on the training day. On test days the chambers had a solid floor and wallpaper. On test day 1, 24 h after the training day, each mouse was returned to the chamber with the new context and exposed to five 30-s tones over 12 min. On test day 2, the mouse was returned to the chamber and exposed to three 30-s tones over 7 min. Beam breaks were measured every second. If there was no new beam break during a 1-s interval, the mouse was considered to be freezing during that interval. The chamber was also equipped with a video camera mounted in the corner for subsequent hand-scored freezing, which was measured as the time during which the mouse exhibited no movement except for breathing. The amount of time spent freezing was expressed as a percentage of total session time.

Cocaine CPP was measured in non-cannulated *Prkcz*^{−/−} and wild-type mice based on the protocol in ref. 15, and in cannulated *Prkcz*^{−/−} mice based on the protocol in ref. 3, but using a two-chambered apparatus (Med Associates Inc.) and 20-min conditioning sessions. Injections of saline or 10 mg kg^{−1} cocaine intraperitoneally (Sigma-Aldrich) were counterbalanced across groups. Cocaine preference index was calculated as the time (seconds) spent in the cocaine-paired chamber on test day minus the time spent in the same chamber before conditioning. Three mice in the non-cannulated group (one wild-type and two *Prkcz*^{−/−} mice) and three *Prkcz*^{−/−} mice (two saline-treated and one ZIP-treated) in the cannulated group did not develop CPP and were therefore excluded from analysis. One outlier in the cannulated group that was identified by a Grubb's test was also removed from the analysis.

Surgery and microinjection. Mice were anaesthetized with ketamine (100 mg kg^{−1} intraperitoneally) and xylazine (7 mg kg^{−1} intraperitoneally) and placed in a digital stereotaxic alignment system (model 1900, David Kopf Instruments). Bilateral guide cannulae (C235GS-5-2.0, 26 gauge, Plastics One) were aimed at the nucleus accumbens (1.40 mm anterior to bregma, ±1.0 mm mediolateral, −3.8 mm ventral from skull surface) and secured with dental cement (DenMat). Mice recovered from surgery for 1 week before the start of experiments. The amount of ZIP peptide (Tocris Bioscience, R&D Systems) in a 1-mg vial was assessed using a guanidine hydrochloride-based Bradford assay (Sigma-Aldrich). The reported peptide purity by Tocris Bioscience closely matched the measured peptide purity. ZIP was dissolved in 0.9% physiological saline, adjusting for peptide purity, to a 10 mM concentration (10 nmol μl^{−1}). Mice were injected with 1 μl of ZIP or saline per side at 0.25 μl min^{−1} using injectors that extended 0.7 mm beyond the guide cannulae. The injectors were left in place for 1 min to allow for diffusion, after which they were removed and the obstructers replaced. The mice were returned to their home cage after injection.

Histological verification of cannulae placements. After completion of the experiment, mice were killed and the brain was removed and placed in 4% paraformaldehyde in 0.1 M phosphate buffer, pH 7.4. The brains were transferred to a 20% sucrose solution in 4% paraformaldehyde for 2 days. Brains were frozen, cut into sections and Nissl stained to verify cannulae placement. Mice with injection sites outside of the nucleus accumbens were excluded from the analysis.

Kinase assay. Kinase activity was measured using the LANCE PKC Assay kit (PerkinElmer Life Sciences). Flag-affinity purified rat PKC-ζ (1 nM), PKM-ζ (0.5 nM) or PKC-ι (2.5 nM) was added to the buffer with 50 nM ULight-PKC peptide substrate, CRFARKGSLRQKNV, (TRF0108-D, PerkinElmer Life Sciences) and increasing concentrations of test compound. The reaction was initiated by adding 2.5 μM ATP and terminated after 60 min by adding 2× Stop Solution/ Detection Mix containing 20 mM EDTA and 4 nM Eu-anti-phospho-PKC (Ala25Ser) (TRF0207-D, PerkinElmer Life Sciences). Increasing concentrations of the PKC peptide substrate (2.5–50 nM) were used to determine K_m values for each atypical PKC. ZIP and scrambled ZIP were obtained from Tocris Bioscience and were dissolved in 0.9% physiological saline, after adjusting for reported peptide purity. Chelerythrine was obtained from Sigma-Aldrich, dissolved in DMSO and assayed using PKM-ζ prepared in the absence of DTT as chelerythrine changed colour and lost all inhibitory activity when 1 mM DTT was present. Phosphorylation was detected using a FlexStation III Microplate Reader in LANCE TR-FRET mode (excitation = 340 nm, emission = 665 nm) and was expressed as relative fluorescence units (RFU). The percentage of inhibition by each test compound was calculated as: (signal without test compound − signal with test compound)/ (signal without test compound − signal in the absence of ATP) × 100. Data were analysed by nonlinear regression and K_i values were calculated by the Cheng–Prusoff equation using Prism 5.0c (GraphPad Software).

23. Chen, J. *et al.* The type 1 equilibrative nucleoside transporter regulates anxiety-like behavior in mice. *Genes Brain Behav.* **6,** 776–783 (2007).
24. Hodge, C. W. *et al.* Decreased anxiety-like behavior, reduced stress hormones, and neurosteroid supersensitivity in mice lacking protein kinase Cε. *J. Clin. Invest.* **110,** 1003–1010 (2002).

LETTER

Oceanographic controls on the diversity and extinction of planktonic foraminifera

Shanan E. Peters[1], Daniel C. Kelly[1] & Andrew J. Fraass[1]†

Understanding the links between long-term biological evolution, the ocean–atmosphere system and plate tectonics is a central goal of Earth science. Although environmental perturbations of many different kinds are known to have affected long-term biological evolution, particularly during major mass extinction events[1,2], the relative importance of physical environmental factors versus biological interactions in governing rates of extinction and origination through geological time remains unknown[2]. Here we use macrostratigraphic data from the Atlantic Ocean basin to show that changes in global species diversity and rates of extinction among planktonic foraminifera have been linked to tectonically and climatically forced changes in ocean circulation and chemistry from the Jurassic period to the present. Transient environmental perturbations, such as those that occurred after the asteroid impact at the end of the Cretaceous period[1] approximately 66 million years ago, and the Eocene/Oligocene greenhouse–icehouse transition[3,4] approximately 34 million years ago, are superimposed on this general long-term relationship. Rates of species origination, by contrast, are not correlated with corresponding macrostratigraphic quantities, indicating that physiochemical changes in the ocean–atmosphere system affect evolution principally by driving the synchronous extinction of lineages that originated owing to more protracted and complex interactions between biological and environmental factors.

Planktonic foraminifera are extant amoeboid protists that inhabit the surface ocean from the tropics to the poles, and their calcium carbonate tests are abundant as microfossils (<1 mm) in marine sediments globally, making them ideal tools for biostratigraphic correlation and reconstructing past environmental conditions[5]. Planktonic foraminifera also constitute a volumetrically significant source of the sediment that is deposited on oceanic crust. Approximately $55 \times 10^6 \, \text{km}^2$ (~67%) of the Atlantic Ocean's sea floor is blanketed by biogenic calcareous ooze[6], and the tests of planktonic foraminifera are an important constituent, contributing up to 80% of the deep-sea sedimentary calcite budget annually[7]. Previous work has attributed macroevolutionary patterns among planktonic foraminifera to several biological characteristics as well as to changes in the physical and chemical state of the ocean. For example, several lineages have independently acquired photosymbionts, an advantageous trait in oligotrophic settings[8]. The delivery of nutrients to the surface ocean and the partitioning of trophic resources between shelf and open ocean settings have also been linked to the abundance and diversification of planktonic foraminifera[9]. However, as highly effective dispersers in one of the most expansive biomes on Earth[10], the degree of partitioning of their ocean habitat through water-column stratification, latitudinal temperature gradients, sea-level fluctuations, and ocean circulation patterns is generally thought to be the most important determinant in the evolution of planktonic foraminifera[11,12].

Here we test the long-standing hypothesis that changes in the physiochemical state of the ocean have exerted a dominant influence on the long-term macroevolution of planktonic foraminifera. To do so,

we compare patterns of taxonomic diversity and rates of extinction and origination to analogous macrostratigraphic quantities that describe spatiotemporal patterns of sedimentation in the deep sea. Macrostratigraphy provides a useful quantitative framework for testing this hypothesis because deep-sea sedimentation is sensitive to many important oceanographic factors, including the location of oceanic-bottom water production and its chemistry, atmospheric CO_2 concentration and the position of the calcite compensation depth (CCD), the routeing and intensity of ocean currents, and the location and extent of nutrient and solute delivery to the surface ocean[13,14]. Thus, spatiotemporal shifts in the location and extent of deep-sea sedimentation provide an integrated signal of physiochemical changes in the ocean–atmosphere system[13–16].

The total number of Atlantic drilling sites that intersect with the sea floor of a given age (Fig. 1; see Methods Summary) increases monotonically from the Jurassic period to the present (Fig. 2a). This long-term increase reflects the opening of the Atlantic Ocean basin during

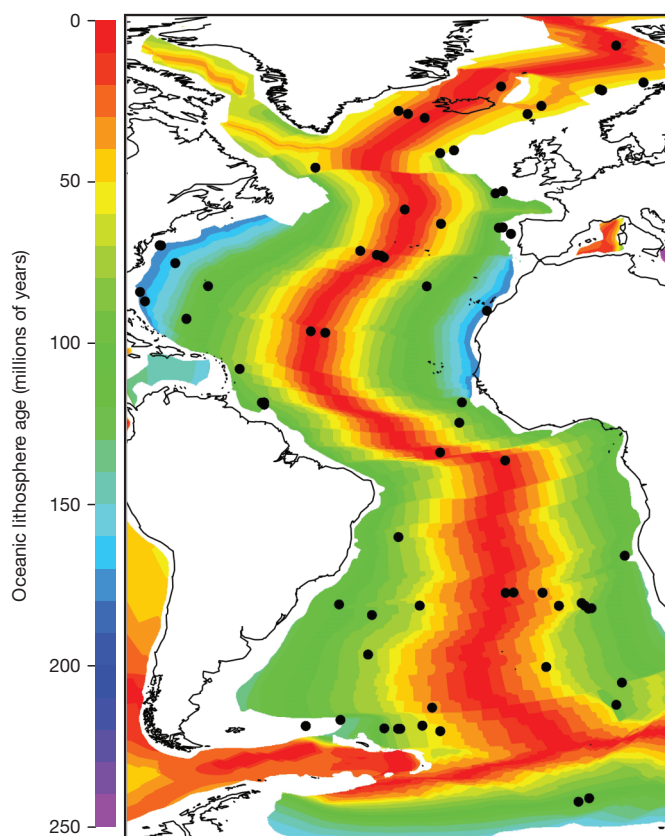**Figure 1 | Atlantic Ocean basin lithosphere ages and locations of 73 sampled drilling sites.** Note that, on average, nearly 500 m of younger sediments overlie oceanic crust at each site (black dots).

[1]Department of Geoscience, University of Wisconsin, Madison, Wisconsin 53706, USA. †Present address: Department of Geoscience, University of Massachusetts-Amherst, Amherst, Massachusetts 01003, USA.
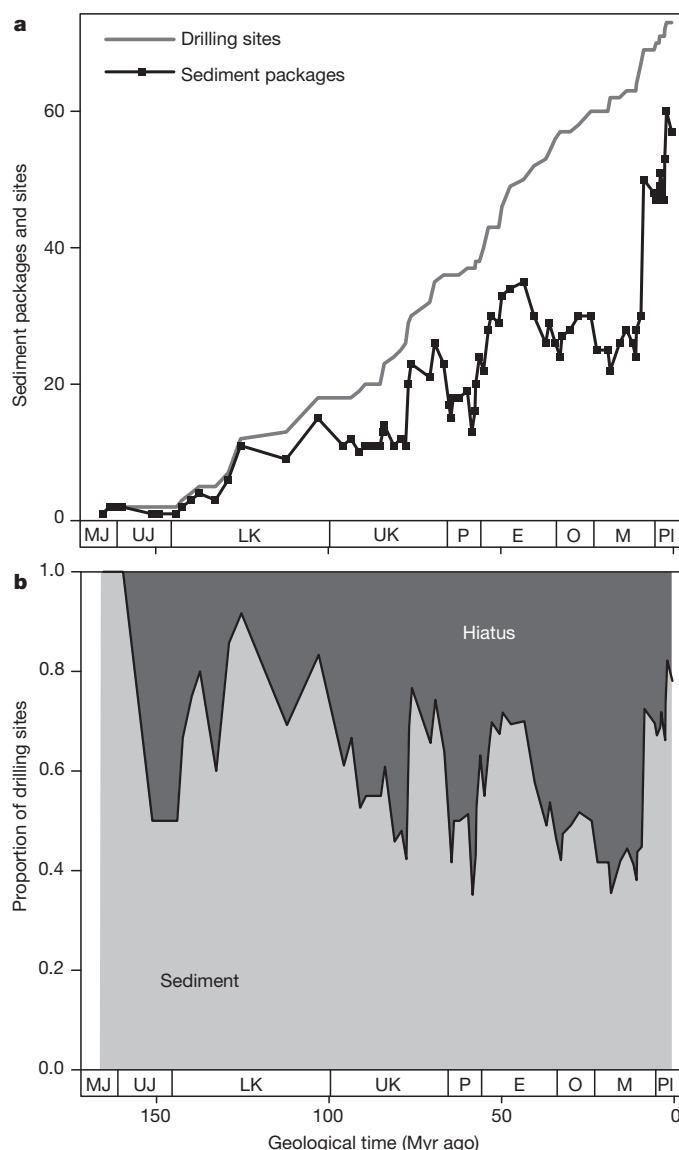
Figure 2 | Macrostratigraphy of the Atlantic Ocean basin. a, Total number of drilling sites intersecting the sea floor and the total number of sediment packages as a function of geologic time. **b,** Proportion of drilling sites with sediment; calculated as the total number of sites with sediment of a given age (black line in **a**) divided by the total number of sites (grey line in **a**). Myr, millions of years. Data are plotted at the age of the base of calcareous nannoplankton zones. MJ, Middle Jurassic; UJ, Upper Jurassic; LK, Lower Cretaceous; UK, Upper Cretaceous; P, Palaeocene; E, Eocene; O, Oligocene; M, Miocene; Pl, Pliocene–Pleistocene.

the Jurassic period and its ongoing expansion through sea-floor spreading. However, the total number of sites that actually preserve sediment of a given age does not increase monotonically over time owing to expansions and contractions in the proportion of sea-floor-preserving sediments. On average, ~60% of the sites in the Atlantic Ocean that could have preserved sediments at any given time actually did, but this proportion varies from less than 40% in the Palaeocene epoch to over 80% at various times during the Mesozoic era and Quaternary period (Fig. 2b).

The decline in the extent of sedimentation in the Atlantic basin from the Lower Cretaceous to the early part of the Upper Cretaceous period, and the subsequent increase to an Upper Cretaceous peak (Fig. 2), reflects shifts in the shelf-to-basin partitioning of carbonate sedimentation that have been linked to a prolonged rise and then fall in global sea level[17], a major change in the global ocean that is exemplified

by the formation and draining of the North American Interior Seaway[18]. It has been postulated that the withdrawal of epicontinental seaways promoted an offshore shift in carbonate deposition during the Upper Cretaceous, with the ancillary effect of depressing the CCD[19]. Individual deep-sea records are consistent with a deepening of the CCD during this time[20], and the basin-wide increase in sedimentation during the Upper Cretaceous documented here reflects this offshore shift in carbonate burial[19].

The end of the Cretaceous period (end-Cretaceous) has long been recognized as a period of abrupt change in the ocean–atmosphere system, much of which is attributable to the Chicxulub asteroid impact[1]. The aftereffects of the end-Cretaceous crisis were prolonged[21], resulting in suppressed deep-sea sedimentation during much of the Palaeocene. This Palaeocene nadir in deep-sea sedimentation, which has been recognized globally[15], ended as sedimentation increased in extent across the Palaeocene/Eocene boundary, eventually reaching a Palaeogene maximum during the early-to-middle Eocene (Fig. 2b), possibly due to enhanced continental weathering in a globally warm climate[3]. Sedimentation in the Atlantic then declined in areal extent during the late Eocene epoch, resulting in a minimum in the early Oligocene epoch. This decline in sedimentation across the Eocene/Oligocene boundary coincides with the largest climate shift in the past 200 million years, a transition characterized by global cooling, expansion of Antarctic ice sheets, fall in sea level, and steepening of meridional temperature gradients[3,4,17].

A protracted Oligocene-to-Miocene interval of relative stasis or decline in the extent of Atlantic sedimentation was terminated by the largest expansion in sedimentation in the basin's history (Fig. 2). Like the Upper Cretaceous, the late Neogene peak in sedimentation is associated with deepening of the CCD in the Atlantic[14]. This event marks the initial development of modern-like thermohaline circulation, which involves routeing of Northern Component Water depleted in respired CO_2 towards the south and a concomitant deepening of the CCD throughout much of the Atlantic basin[22].

Macrostratigraphy quantifies spatiotemporal changes in deep-sea sedimentation that are attributable to physiochemical changes in the ocean–atmosphere system, many of which have been recognized[14–16] and identified within individual deep-sea records[23]. The independently derived global macroevolutionary history of planktonic foraminifera bears many quantitative similarities to the macrostratigraphy of the Atlantic (Fig. 3). Similarly to the number of sediment packages, there
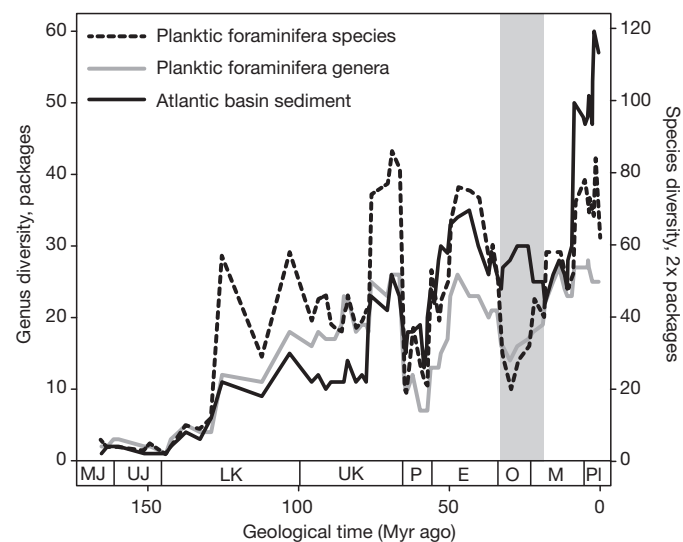

Figure 3 | Macrostratigraphy of the Atlantic Ocean basin and global range through diversity of planktonic foraminifera. De-trended time series are positively correlated (Table 1). A transient decoupling and shift to negative correlation occurred during the Oligocene–early Miocene (shaded interval). See Fig. 2 legend for abbreviations.

Table 1 | Macrostratigraphy–macroevolution correlations

	Species		Genera	
	r_s	P	r_s	P
Diversity–packages	0.526	$<1.0 \times 10^{-5}$	0.442	0.00011
Extinction–truncation	0.491	0.00016	0.524	0.00005
Origination–initiation	0.250	0.031	0.261	0.030

Spearman rank-order correlation coefficients (r_s) and P values between de-trended (using first differences) time series of global macroevolutionary quantities for planktonic foraminifera and corresponding macrostratigraphic quantities[30] for the Atlantic Ocean basin. To avoid edge effects and time intervals with unconstrained rates, genus-level comparisons range from NN20 to NJ12 (0.3 to 165.4 million years ago), species-level comparisons from NN20 to CC2 (0.3 to 142.4 million years ago). P values were derived from 100,000 bootstrap randomizations.

are three peaks in the global species and generic diversity of planktonic foraminifera separated by two minima, one in the Palaeocene and one in the Oligocene. The first differences in diversity, which emphasize interval-to-interval changes rather than longer-term trends, are significantly positively correlated with changes in the number of sediment packages (Table 1, correlations remain significant ($P \le 0.006$) when the effect of interval duration is removed using partial correlation; Supplementary Table 1). However, the positive correlation between sediment packages and diversity is negative from the Oligocene to early Miocene (Fig. 3, shaded interval), weakening the overall correlations reported in Table 1. The first differences in per capita rates of extinction are also significantly positively correlated with rates of sediment-package truncation (Table 1, extinction correlations remain significant ($P \le 0.022$) after removing the effects of interval duration; Supplementary Table 1), with the only outlier being a much larger than expected increase in species extinction in the aftermath of the end-Cretaceous asteroid impact[1]. Correlations between rates of origination and sediment-package initiation are, by contrast, weaker (Table 1) and not significant after adjusting for interval duration ($P \ge 0.09$; Supplementary Table 1).

Given that there are errors associated with both the macroevolutionary and macrostratigraphic data (see Methods), the presence of statistically significant correlations is remarkable and implies that the records are related. Comparable correlations between the sedimentary and fossil records in shallow marine and continental shelf environments have been interpreted both as a preservation-induced sampling bias in fossil data and as a shared, common signal that is induced by environmental changes associated with expansions and contractions in epicontinental seas[24].

The correlations between global macroevolutionary patterns in planktonic foraminifera and the macrostratigraphy of the Atlantic Ocean basin (Table 1) are unlikely to be due to a fossil preservation bias for several reasons. First, planktonic foraminifera are widespread geographically and only a few sites (approximately five to seven) distributed across latitudes are required to recover all extant species (see Methods). Yet, drilling sites yielding foraminifera number in the thousands globally, with almost all time intervals since the Jurassic represented by many more than seven sites. Second, the ability of Atlantic macrostratigraphy to predict global extinction better than origination (Table 1) is not consistent with the hypothesis that preservation biases have artificially clustered the first and last appearances of species at hiatuses in the sedimentary record (see Methods). Third, there is a reversal in the sign of the correlation between diversity and the number of sedimentary packages during part of the Oligocene–Miocene periods (Fig. 3), which is inconsistent with a sediment-quantity-controlled sampling bias. Fourth, the overall apparent increase in the extent of deep-sea sedimentation towards the present (Fig. 2a), a global signal attributable to sea-floor spreading and oceanic lithosphere subduction, is not mirrored by an increase in diversity, as predicted by the hypothesis that fossil preservation in sediments is driving the correlations. Last, adjusting face-value macroevolutionary rates to account for variably incomplete sampling explicitly does not qualitatively change our results. See Methods for further discussion of this optimization approach and other sampling bias related issues.

For all of these reasons, our results are best interpreted as evidence for a causal link between the processes that drive patterns of deep-sea sedimentation and the global diversity and extinction of planktonic foraminifera. Under this hypothesis, the transient decoupling of diversity and sedimentary packages that occurs at the Eocene/Oligocene boundary is attributed to a lagged macroevolutionary response to the most profound climate change of the past 200 million years[12]. Key aspects of the calcareous plankton response to the Eocene/Oligocene climatic transition are the selective extinction of warm-water taxa, equator-ward migrations of high-latitude taxa, and reduced diversities in subpolar-to-polar bioprovinces[25]. Collectively, these lines of evidence suggest that global cooling and enhanced ocean mixing reduced thermal stratification in high latitudes, thereby reducing the amount of ecospace available for segregating planktonic foraminiferal species on the basis of depth ecology[11,12]. The Oligocene increase in the extent of sedimentation, signalled by a modest upturn in sediment packages (Fig. 2), has been previously noted and attributed to a deepening of the CCD in the wake of the Eocene/Oligocene transition[26]. Thus, the reorganization of the ocean–climate system during the Eocene/Oligocene transition probably served as a common cause, driving both patterns of deep-sea sedimentation and diversity.

The ability of macrostratigraphy to predict the extinction of planktonic foraminifera better than origination is difficult to explain under the hypothesis of sampling bias, but it is readily interpretable as a biological signal. The discovery that most extant morphospecies of planktonic foraminifera consist of multiple genotypes with an extraordinary capacity to disperse, challenges the view that tectonic and hydrographic barriers facilitate speciation by isolating reproductive populations[27,28]. Instead, the genetic and zoogeographic evidence suggest that the evolution of new planktonic foraminiferal species may be limited by their ability to maintain populations in areas where unfavourable ecological conditions exist rather than by an inability to disperse beyond tectonic and hydrographic barriers[10]. An important corollary of this finding is that biotic processes such as competitive exclusion between sibling species, establishment of symbiotic relationships, differences in trophic behaviour, and divergent reproductive strategies may have a more prominent role in the origination of new species than the tectonic and climatic perturbations that consistently drive synchronous changes in the physiochemical structure of the ocean, the distribution and extent of deep-sea sedimentation, and species extinction.

METHODS SUMMARY

Data for aspects of this analysis were derived from Macrostrat (http://macrostrat.org). We compiled data on the lithology, thickness and age of sediments recovered at each of 73 offshore sites in the Atlantic Ocean (Fig. 1). Although decades of offshore drilling have resulted in globally distributed data, the Atlantic is the only basin where geographical coverage is sufficient to provide estimates of basin-scale patterns of sedimentation. Only those sites that were cored continuously from the sea floor down to or near the basaltic ocean crust were included to avoid characterizing the idiosyncratic sampling of deep-sea sediments, as opposed to their actual spatiotemporal distribution. The average modern water depth of all 73 Atlantic sites is 3.28 km, which is slightly less than the 3.92 km average depth of the Atlantic basin, indicating a modest preference for drilling on oceanic rises and in shallower water locations.

Hiatuses, which define the boundaries between otherwise continuously accumulated sediment packages, form the basis of macrostratigraphy[29,30], and were identified at each site wherever one or more calcareous nannofossil biozones (median duration 1.8 million years) were missing or not represented by a measurable thickness of sediment. A total of 249 hiatus-bound sediment packages composed of 3,692 lithologic units were recognized across all 73 Atlantic sites. Results are similar if 59 additional sites from other ocean basins, cored appropriately for macrostratigraphy, are included (see Supplementary Information).

The macroevolutionary history of planktonic foraminifera was determined by independently compiling the global stratigraphic ranges of 671 species documented in published taxonomic atlases and online resources. First and last appearance datums for planktonic foraminiferal species were binned into the same calcareous nannoplankton biozones used to recognize deep-sea hiatuses. Rates of species

origination and extinction, and rates of sediment package initiation and truncation, were both calculated using per capita, per interval turnover rates[30].

Full Methods and any associated references are available in the online version of the paper.

1. Schulte, P. et al. The Chicxulub asteroid impact and mass extinction at the Cretaceous-Paleogene boundary. Science **327,** 1214–1218 (2010).
2. Benton, M. J. The Red Queen and the Court Jester: species diversity and the role of biotic and abiotic factors through time. Science **323,** 728–732 (2009).
3. Zachos, J. C., Pagani, M., Sloan, L., Thomas, E. & Billups, K. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science **292,** 686–693 (2001).
4. Liu, Z. et al. Global cooling during the Eocene-Oligocene climate transition. Science **323,** 1187–1190 (2009).
5. Vincent, E. & Berger, W. H. in The Oceanic Lithosphere (ed. Emiliani, C.) 1025–1119 (Wiley, 1981).
6. Sverdrup, H. U., Johnson, M. W. & Fleming, R. H. The Oceans: Their Physics, Chemistry, and General Biology (Prentice-Hall, 1942).
7. Schiebel, R. Planktonic foraminiferal sedimentation and the marine calcite budget. Glob. Biogeochem. Cycles **16,** 1065 (2002).
8. D'Hondt, S. & Zachos, J. C. Cretaceous foraminifera and the evolutionary history of planktonic photosymbiosis. Paleobiology **24,** 512–523 (1998).
9. Hallock, P., Premoli-Silva, I. & Boersma, A. Similarities between planktonic and larger foraminiferal evolutionary trends through Paleogene paleoceanographic changes. Palaeogeogr. Palaeoclimatol. Palaeoecol. **83,** 49–64 (1991).
10. Norris, R. D. Pelagic species diversity, biogeography, and evolution. Paleobiology **26,** 236–258 (2000).
11. Lipps, J. H. Plankton evolution. Evolution **24,** 1–22 (1970).
12. Cifelli, R. Radiation of Cenozoic planktonic foraminifera. Syst. Zool. **18,** 154–168 (1969).
13. Berggren, W. A. & Hollister, C. D. Plate tectonics and paleocirculation—commotion in the ocean. Tectonophysics **38,** 11–48 (1977).
14. Van Andel, T. H. Mesozoic/Cenozoic calcite compensation depth and the global distribution of calcareous sediments. Earth Planet. Sci. Lett. **26,** 187–194 (1975).
15. Moore, T. C. Jr, van Andel, T. H., Sancetta, C. & Pisias, N. Cenozoic hiatuses in pelagic sediments: marine plankton and sediments. Micropaleontol. Sp. Pub. **3,** 113–138 (1978).
16. Keller, G. et al. Global distribution of late Paleogene hiatuses. Geology **15,** 199–203 (1987).
17. Miller, K. G. et al. The Phanerozoic record of global sea-level change. Science **310,** 1293–1298 (2005).
18. Kauffman, E. G. & Caldwell, W. G. E. Evolution of the Western Interior Basin 1–30 (Geological Association of Canada Special Paper 39, 1993).
19. Milliman, J. D. in Recent Sedimentary Carbonates (eds Milliman, J. D., Müller, G. & Förstner, U.) Ch. 8 (Springer, 1974).
20. Thierstein, H. R. in Deep Sea Drilling Results in the Atlantic Ocean: Continental Margins and Paleoenvironment (eds Talwani, M., Hay, W. W. & Ryan, W. B. F.) 249–274 (American Geophysical Union, 1979).
21. D'Hondt, S., Donaghay, P., Zachos, J. C., Luttenberg, D. & Lindinger, M. Organic carbon fluxes and ecological recovery from the Cretaceous-Tertiary mass extinction. Science **282,** 276–279 (1998).
22. Lyle, M., Dadey, K. A. & Farrell, J. W. The late Miocene (11–8 Ma) eastern Pacific carbonate crash: evidence for reorganization of deep-water circulation by the closure of the Panama Gateway. Proc. Ocean Drill. Program Sci. Results **138,** 821–838 (1995).
23. Norris, R. D., Klaus, A. & Kroon, D. in Western North Atlantic Paleogene and Cretaceous Paleoceanography (eds Kroon, D., Norris, R. D. & Klaus, A.) 23–48 (Geological Society of London, 2001).
24. Hannisdal, B. & Peters, S. E. Phanerozoic Earth system evolution and marine biodiversity. Science **334,** 1121–1124 (2011).
25. Haq, B. U., Premoli-Silva, I. & Lohmann, G. P. Calcareous plankton paleobiogeographic evidence for major climatic fluctuations in the Early Cenozoic Atlantic Ocean. J. Geophys. Res. **82,** 3861–3876 (1977).
26. Thunell, R. C. & Corliss, B. H. in Terminal Eocene Events (eds Pomerol, C. & Premoli-Silva, I.) 363–380 (Elsevier, 1986).
27. Darling, K. F., Wade, C. M., Kroon, D., Brown, L. & Bijma, J. The diversity and distribution of modern planktonic foraminiferal SSU rRNA genotypes and their potential as tracers of present and past ocean circulation. Paleoceanography **14,** 3–12 (1999).
28. Darling, K. F. et al. Molecular evidence for genetic mixing of Arctic and Antarctic subpolar populations of planktonic foraminifers. Nature **405,** 43–47 (2000).
29. Peters, S. E. Macrostratigraphy of North America. J. Geol. **114,** 391–412 (2006).
30. Hannisdal, B. & Peters, S. E. On the relationship between macrostratigraphy and geological processes: quantitative information capture and sampling robustness. J. Geol. **118,** 111–130 (2010).

Supplementary Information is available in the online version of the paper.

Author Contributions All authors compiled Macrostrat data and contributed to development of the manuscript; D.C.K. and A.J.F. compiled the planktonic foraminifera species range data.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.E.P. (peters@geology.wisc.edu).

METHODS

Planktonic foraminifera data. Species stratigraphic ranges were compiled from published atlases and online sources to ensure coherent, consistent and robust taxonomic statuses for all species, as well as community-consensus-based stratigraphic ranges. Global stratigraphic ranges for Jurassic and Cretaceous species were retrieved from the Mesozoic Planktonic Foraminiferal Taxonomic Dictionary[31] and PLANKRAGE[32] circa 2008. Ranges of Palaeogene species were largely adopted from refs 33, 34. The primary reference for Neogene species was ref. 35, with supplementary data from ref. 36. Other references used to compile the data set included refs 37, 38. Priority for both taxonomic and range data were given to the most recently published work. Generic ranges were compiled from the oldest first appearance datum and the youngest last appearance datum for all the species within a given genus. For this analysis, all first and last appearances were placed into temporal bins defined by calcareous nannoplankton biozones[39].

Factors affecting the precision of the macroevolutionary quantities presented here can be separated into two different types: (1) taxonomic practice; and (2) correlation. Errors of taxonomic practice principally involve species concepts and the recognition and subdivision of lineages into discrete species- and genus-level units (note that we present results for both and they are comparable). These taxonomic issues have been discussed in the literature[40,41], and although such factors are likely to have added variance to the macroevolutionary quantities studied here, we expect them to be random effects with respect to macroevolutionary patterns, as has been documented in the case of marine invertebrates in marine shelf environments[41]. Errors of correlation principally involve the accurate assignment of ages to the fossil occurrences that define the first and last appearances of lineages, and the correlation (if necessary) of those ages to the calcareous nannofossil zones used here. These too have been shown to be random effects in the case of marine invertebrates[41].

Macrostratigraphic data. Data for aspects of this analysis were derived from Macrostrat (http://macrostrat.org). Macrostratigraphy is based on the times of initiation and truncation of sediment packages that are continuous at a specified scale of temporal resolution and that are identified at multiple locations distributed across a geographical region of interest[29,30,42]. In the case of this analysis, temporal resolution was defined as calcareous nannofossil zones and the geographical region of interest was the Atlantic basin, which has been continuously expanding in area since its initial opening in the Triassic–Jurassic[43]. Thus, the hiatuses that are used to recognize package boundaries are here defined as the point in a core where one or more calcareous nannofossil zones are not represented by a significant thickness of sediment (see Supplementary Information for an example). In most cases, the biostratigraphy of the core is well constrained and the hiatuses are clearly defined, but in some cases hiatuses are marked by relatively thin intervals of terrigenous clay that internally lack the calcareous microfossils that are most widely used to obtain age control. In some clay-rich intervals, siliceous and organic walled microfossils can be used to obtain biostratigraphic age control, but some clays are barren even of these fossils. Thus, the age of clay in deep-sea cores is often poorly constrained, but it is likely that most of the clay-rich intervals accumulated very slowly as a result of wind-blown continental input and no biogenic input. Thus, the effective rates of sedimentation are so low as to constitute a significant hiatus relative to the more abundant calcareous and siliceous sediments that characterize most of the volume of sediment in the deep sea.

Factors affecting the precision and accuracy of the macrostratigraphic quantities presented here are of two types. The first involves the spatial distribution and number of control points that are used to characterize the sedimentary record (Fig. 1). This property of our data has been dictated by the nature and practice of offshore drilling. There are many hundreds of offshore drilling sites in the Atlantic, but only those sites that were cored continuously from the sea floor down to or very near the basement are appropriate for inclusion in macrostratigraphic analyses. All appropriate Atlantic sites have been included here, so our data are as complete as possible in spatial coverage at this time. The fact that the curve describing the total number of sites (Fig. 2a, grey line) has a nearly constant slope suggests that our coverage relative to the age of the underlying lithosphere is satisfactory. This is not the case in other ocean basins, which is why we have limited our macrostratigraphic analysis to the Atlantic. It is expected that should comprehensive global coverage of macrostratigraphic data become available, the correlations we report here would strengthen.

The second source of error involves the recognition of hiatuses and the correlation of their beginning and end points in cores. In some cases, the biozonation of sediments above and/or below a hiatus is not definitive. For example, an assemblage might be consistent with Eocene Nannofossil Zones NP14 and NP15. In such cases, we used the most conservative age with respect to the adjacent hiatus. That is, if sediments zoned as NP14/15 were above a hiatus, the sediments would be assigned a maximum initiation age of NP14. If sediments zoned as NP14/15 were

below a hiatus, they would be assigned a minimum truncation age of NP15, thereby minimizing the hiatus duration in both directions.

The effects of variable and incomplete sampling. An often-asked question regarding palaeodiversity curves, such as the one tabulated here from the literature, pertains to how completely the available fossil record, and the sampling thereof, captures true biological diversity and the times of actual origination and extinction of lineages. This is a reasonable question in the present study, given that all of our planktonic foraminiferal diversity estimates are ultimately based on samples and because the pelagic realm they inhabit is one of the largest biomes on Earth. However, it has long been recognized that the risk of failing to sample most planktonic foraminiferal species due to inadequate spatial coverage is greatly reduced by the expansive biogeographic ranges of the constituent species[44] and by their large numbers in even small volumes of sediment. This supposition is supported by studies of both living samples collected in plankton tows and by samples of death assemblages preserved in surface sediments on the sea floor[45–47]. All of these investigations have shown that modern planktonic foraminiferal species occur in five major faunal provinces, which are arrayed along meridional temperature gradients. These five provinces are fully encompassed by the north–south-oriented Atlantic Ocean basin[48,49]. Thus, only a small volume of sediment from a few locations (approximately five) strategically positioned along a latitudinal transect recovers nearly all modern planktonic foraminiferal species. The number of offshore drilling sites globally, by contrast, is in the thousands, with hundreds of kilometres of core recovered and examined. Marine sections located on continents and exposed in outcrops also contribute to the sampling of planktonic foraminifera. Thus, all geological time intervals relevant to the history of planktonic foraminifera, at least from the Cretaceous to the present, are represented by multiple sites distributed across the latitudes and basins required to capture modern species diversity. For these reasons, we consider the diversity curves presented here, which are based on authoritative summaries compiled on a per taxon basis, not on stratigraphic ranges derived *de novo* on a per sample basis, to capture secular variation in planktonic foraminiferal diversity accurately, at least since the late Jurassic. Similarly, we believe that the variability in deep-sea sedimentation documented here (Fig. 2), although reflecting actual changes in the extent of sediment preservation on the sea floor in the Atlantic, is not a relevant measure of the sampling intensity of planktonic foraminifera globally.

Although there is abundant evidence to suggest that only a small number of sites are required to capture global diversity of planktonic foraminifera species, and therefore that their global sampling is quite complete, at least since the late Jurassic, one way to quantify the effect of sampling is to use an occurrence-based database of planktonic foraminifera and to determine the minimum number of sites that are required to recover all species within that database. To this end, we used the CLIMAP data set[50]. One of the objectives of the CLIMAP project was to compile core-top data from a large number of drilling sites, and many of these core-top data are associated with planktonic foraminiferal species lists. However, the CLIMAP planktonic foraminiferal data are not comprehensive because of the scientific focus of the CLIMAP project. Only the 39 species deemed informative in palaeoclimate reconstruction were tracked in this survey, which is approximately 55% of the total number of planktonic foraminiferal species that were probably encountered in the samples. Nevertheless, we determined the minimum number of sites required to recover all 39 species present in the CLIMAP data set by randomly resampling the CLIMAP planktonic-foraminifera-bearing sites and then calculating diversity for a subset of the available sites. We find that to capture all species within the CLIMAP data set, a minimum number of six sites must be included and the average sampling curve increases very steeply (see Supplementary Information).

Another way to account for variable and incomplete sampling is to derive estimates for sampling probabilities and then to correct the empirical rates of origination and extinction for variable and incomplete preservation. This was done using a rate optimization approach that simultaneously finds the best-fit true rates of extinction, origination and preservation, given only the empirical taxonomic survivorship matrices[51]. This rate optimization routine was executed in July 2012 (M. Foote, personal communication) using species-level survivorship matrices binned temporally into calcareous nannofossil zones[39]. A total of 77 optimizations using the pulsed turnover model[51] converged on stable solutions; the continuous time model produced no usable results due to lack of convergence. The mean values for 77 independently fitted rates in each time interval were calculated to produce a single time series describing the average best-fit optimized rates. See ref. 51 for more information about the rate-fitting procedure and its sensitivity to a variety of empirical and simulated macroevolutionary and preservation scenarios.

The first differences in the mean preservation-corrected species extinction rates are strongly positively correlated with face-value extinction rates ($r_s = 0.808$; $P = 2.2 \times 10^{-16}$). Mean optimized origination rates, by contrast, are less strongly positively correlated with face-value rates ($r_s = 0.547$; $P = 2.1 \times 10^{-6}$). The higher

correlation between optimized extinction and face-value extinction indicates that empirical rates of extinction among planktonic foraminifera have a higher fidelity in the face-value fossil record than rates of origination (or that origination violates the pulsed model more than extinction), a result that is consistent with similar analyses for marine animals in continental shelf environments[52,53]. Optimized rates of extinction remain more strongly positively correlated with package truncation rates ($r_s = 0.392$; $P = 9.3 \times 10^{-4}$) than optimized rates of origination are with package initiation rates ($r_s = 0.220$; $P = 0.072$). Interestingly, optimized rates of preservation are quite high (mean 0.70, median 0.81) and mean preservation rates are positively correlated with the number of sediment packages ($r_s = 0.406$; $P = 5.9 \times 10^{-4}$). However, the correlation between fitted preservation rates and the number of sediment packages drops markedly and becomes non-significant when the effect of interval duration is removed by partial correlation ($r_{(s \times \Delta t)} = 0.151$; $P = 0.134$). This suggests, not surprisingly, that there is little or no relationship between the global sampling of planktonic foraminifera, as estimated by Foote's rate-fitting algorithms, and the total number of sediment packages in the Atlantic Ocean basin. However, removing the effects of variable interval duration does not lower the correlation between optimized rates of extinction and origination, and the corresponding macrostratigraphic quantities, by as much ($r_{(s \times \Delta t)} \geq 0.220$; $P \leq 0.07$), indicating that a signal remains that cannot be attributed to variable interval duration or preservation.

The global sampling of planktonic foraminifera, which served as the basis of the taxonomic references that were used to determine global stratigraphic ranges on a per lineage basis, is exceedingly comprehensive in comparison to the biogeographic distribution of species. We find that only a few sites (six in the case of CLIMAP data) distributed across latitudinal gradients are required to capture all modern species. Moreover, adjusting rates of extinction and origination using Foote's rate optimization approach, although potentially adding noise to the data if the assumption of pulsed turnover is not always valid[51], does not substantively change the results. Thus, extinction may in fact have higher fidelity in the face-value fossil record than origination, but it remains better predicted by macrostratigraphic quantities than origination. This asymmetry is not expected with a simple notion of sampling bias, which suggests that the incomplete sampling of taxon stratigraphic ranges artificially clusters first and last appearances at hiatuses in the sedimentary record[54–56], although the scaling of these expectations, which were developed for local sections in continental basins, up to the Atlantic basin is not straightforward. Nevertheless, overall, our results suggest that macrostratigraphy and species extinction share a common signal that is driven by tectonics and the evolution of the ocean–atmosphere system.

31. Huber, B. T. *et al.* Mesozoic Planktonic Foraminiferal Taxonomic Dictionary. http://chronos.org (2006).
32. Stewart, D. R. M. & Pearson, P. N. PLANKRANGE: A Database of Planktonic Foraminiferal Ranges. http://palaeo.gly.bris.ac.uk/Data/plankrange.html (2000).
33. Olsson, R. K. *et al. Atlas of Paleocene Planktonic Foraminifera* (Smithsonian Contributions to Paleobiology, 1999).
34. Pearson, P. N. *et al. Atlas of Eocene Planktonic Foraminifera* (Cushman Foundation, 2006).
35. Kennett, J. P. & Srinivasan, M. S. *Neogene Planktonic Foraminifera* (Hutchinson Ross, 1983).
36. Saito, T., Thompson, P. R. & Breger, D. *Systematic Index of Recent and Pleistocene Planktonic Foraminifera* (Univ. of Tokyo Press, 1981).
37. Bolli, H. M., Saunders, J. B. & Perch-Nielsen, K. *Plankton Stratigraphy* (Cambridge Univ. Press, 1989).
38. Georgescu, M. D. A new planktonic foraminifer (family Hedbergellidae Loeblich and Tappan, 1961) from the Lower Campanian sediments of the Falkland Plateau, South Atlantic Ocean (DSDP Site 511). *J. Foraminiferal Res.* **38,** 157–161 (2008).
39. Ogg, J. G., Agterberg, F. P. & Gradstein, F. M. in *A Geologic Time Scale* (eds Gradstein, F. M., Ogg, J. G. & Smith, A. G.) 344–383 (Cambridge Univ. Press, 2004).
40. Purvis, A. Phylogenetic approaches to the study of extinction. *Ecol. Evol. Syst.* **39,** 301–319 (2008).
41. Adrain, J. & Westrop, S. R. An empirical assessment of taxic paleobiology. *Science* **289,** 110–112 (2000).
42. Peters, S. E. Genus extinction, origination, and the durations of sedimentary hiatuses. *Paleobiology* **32,** 387–407 (2006).
43. Müller, R. D., Sdrolias, M., Gaina, C. & Roest, W. R. Age, spreading rates and spreading symmetry of the world's ocean crust. *Geochem. Geophys. Geosyst.* **9,** Q04006 (2008).
44. Prothero, D. R. & Lazarus, D. B. Planktonic microfossils and the recognition of ancestors. *Syst. Zool.* **29,** 119–129 (1980).
45. Parker, F. L. Planktonic foraminiferal species in Pacific sediments. *Micropaleontology* **8,** 219–254 (1962).
46. Bé, A. W. H. & Tolderlund, D. S. in *The Micropaleontology of Oceans* (eds Funnell, B. M & Riedel, W. R.) 1025–1119 (Cambridge Univ. Press, 1971).
47. CLIMAP Project Team Members. The surface of the ice-age Earth. *Science* **191,** 1131–1137 (1976).
48. Bé, A. W. H. in *Oceanic Micropaleontology* (ed. Ramsay, A. T. S.) Vol. 1 1–100 (Academic, 1977).
49. Hemleben, C., Spindler, M. & Anderson, O. R. *Modern Planktonic Foraminifera* (Springer, 1989).
50. CLIMAP Project Members. CLIMAP 18K Database (NOAA/NGDC Paleoclimatology Program, 1994).
51. Foote, M. Origination and extinction through the Phanerozoic: a new approach. *J. Geol.* **111,** 125–148 (2003).
52. Foote, M. Pulsed origination and extinction in the marine realm. *Paleobiology* **31,** 6–20 (2005).
53. Peters, S. E. Geologic constraints on the macroevolutionary history of marine animals. *Proc. Natl Acad. Sci. USA* **102,** 12326–12331 (2005).
54. Holland, S. M. The stratigraphic distribution of fossils. *Paleobiology* **21,** 92–109 (1995).
55. Holland, S. M. The quality of the fossil record: a sequence stratigraphic perspective. *Paleobiology* **26,** 148–168 (2000).
56. Peters, S. E. & Heim, N. A. Stratigraphic distribution of marine fossils in North America. *Geology* **39,** 259–262 (2010).

LETTER

Discrete genetic modules are responsible for complex burrow evolution in *Peromyscus* mice

Jesse N. Weber[1]†, Brant K. Peterson[1,2] & Hopi E. Hoekstra[1,2]

Relative to morphological traits, we know little about how genetics influence the evolution of complex behavioural differences in nature[1]. It is unclear how the environment influences natural variation in heritable behaviour[2], and whether complex behavioural differences evolve through few genetic changes, each affecting many aspects of behaviour, or through the accumulation of several genetic changes that, when combined, give rise to behavioural complexity[3]. Here we show that in nature, oldfield mice (*Peromyscus polionotus*) build complex burrows with long entrance and escape tunnels, and that burrow length is consistent across populations, although burrow depth varies with soil composition. This burrow architecture is in contrast with the small, simple burrows of its sister species, deer mice (*P. maniculatus*). When investigated under laboratory conditions, both species recapitulate their natural burrowing behaviour. Genetic crosses between the two species reveal that the derived burrows of oldfield mice are dominant and evolved through the addition of multiple genetic changes. In burrows built by first-generation backcross mice, entrance-tunnel length and the presence of an escape tunnel can be uncoupled, suggesting that these traits are modular. Quantitative trait locus analysis also indicates that tunnel length segregates as a complex trait, affected by at least three independent genetic regions, whereas the presence of an escape tunnel is associated with only a single locus. Together, these results suggest that complex behaviours—in this case, a classic 'extended phenotype'[4]—can evolve through multiple genetic changes each affecting distinct behaviour modules.

Animal architectures, such as beehives, bird nests, spider webs, termite mounds and rodent burrows, are remarkably diverse traits that can evolve through natural selection. Despite their great diversity, these extended phenotypes have similarities: they seem to be constructed through largely unlearned motor patterns; they are often consistent within a species (or population); and, when architectures differ, these differences reflect important fitness-related functions in the wild[5]. Thus, genetic changes are predicted to contribute to the evolution of different architectures, even between closely related species; however, biologists have long questioned how genetic changes can lead to the evolution of distinct behaviours[6]. Because animal architectures can be precisely measured—similarly to morphological traits, which have been successfully dissected genetically[7,8]—these are excellent traits for the genetic analysis of behaviour.

Many species of *Peromyscus* mice build burrows, which were initially described by natural historians working in the field[9–14]. However, these burrowing behaviours can also be studied in the laboratory[15,16], and our previous work showed that burrow differences among *Peromyscus* mice have a strong genetic component[17]. In particular, *P. polionotus* and *P. maniculatus* are recently diverged and interfertile sister species that build distinct burrows. *P. polionotus* is an open-field specialist and is restricted to the southeastern United States (Fig. 1a and Supplementary Fig. 1a), whereas *P. maniculatus* is a generalist, which inhabits prairie and forest habitats across much of North

America (Supplementary Fig. 1c). Whereas most *Peromyscus* species build small, single-tunnel burrows (Supplementary Fig. 1d) or no burrows at all, *P. polionotus* construct unique burrows characterized by a long entrance tunnel that leads into a nest cavity, and a secondary tunnel that emanates from the nest and terminates just below the soil surface (Fig. 1b and Supplementary Fig. 1b)[9–11]. This secondary tunnel may serve several functions[11], most notably its use as an escape tunnel when predators (for example, snakes) invade the entrance tunnel (see Supplementary Video). Overall, the complex burrow architecture in *P. polionotus* is derived[17], and probably associated with adaptation to living in an exposed, open habitat[15].

To examine natural variation in burrow complexity, we measured burrow dimensions and soil composition (that is, percentage silt, sand and clay) across the range of *P. polionotus* (Fig. 1a and Supplementary
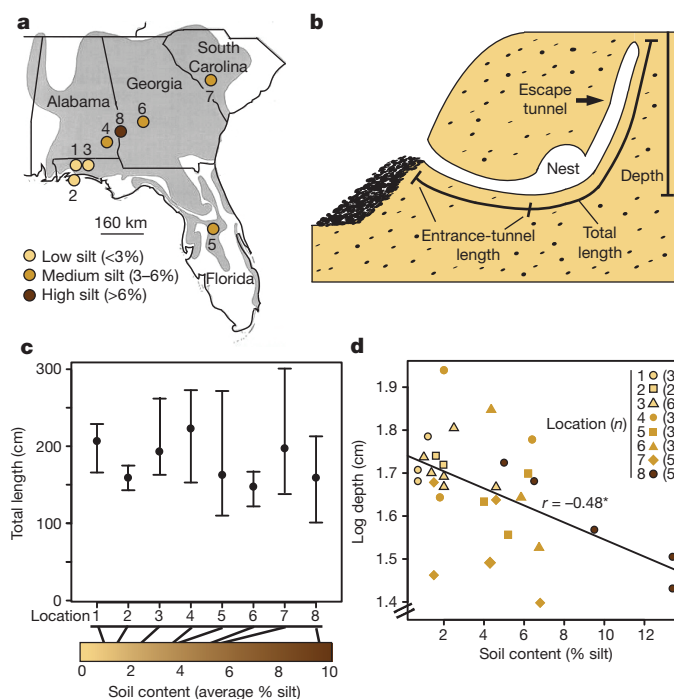

Figure 1 | Natural variation in *P. polionotus* burrows. a, Sampling of burrows at eight sites in the southeastern United States from across the range of *P. polionotus* (grey area). Average percentage of soil silt at each sampling site is provided. **b**, Diagram of a typical *P. polionotus* burrow showing the measures for entrance-tunnel length, total length and burrow depth, as well as a typical escape tunnel. **c**, Variation in total burrow length among sites (mean ± range; from **a**), which are ordered by increasing percentage of silt (left to right). **d**, Correlation between silt composition of soil and burrow depth (asterisk indicates Spearman correlation, rho = −0.48, S = 6632, P = 0.01). Each point represents a burrow, and shapes represent the eight different sampling sites (from **a**). The number of burrows measured at each site is shown in parentheses.

[1]Department of Organismic & Evolutionary Biology, Museum of Comparative Zoology, 26 Oxford Street, Cambridge, Massachusetts 02138, USA. [2]Department of Molecular & Cellular Biology, Center for Brain Science, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA. †Present address: Section of Integrative Biology, One University Station, University of Texas Austin, Texas 78712, USA.

Table 1; see Methods for details). Burrows generally had two long tunnels—an entrance and an escape tunnel (Fig. 1b)—and a mean total length of 181 cm (standard deviation (s.d.) of ± 53). Of the three soil variables measured, the percentage of silt was the most likely to influence soil compaction and thereby influence burrow dimensions[18]. Despite large differences in the percentage of silt among locations (Fig. 1c; Kruskal–Wallis (K–W) test, chi-squared = 19.42, degrees of freedom (df) = 7, $P = 0.01$), burrow length did not differ significantly among locations (K–W test, chi-squared = 6.59, df = 7, $P = 0.47$), nor did burrow length correlate with the percentage of silt measured at each burrow (Spearman test, rho = −0.16, $S = 5,226$, $P = 0.39$). In contrast, there was a significant, negative correlation between the percentage of silt in the soil at each burrow and burrow depth (Fig. 1d; Spearman test, rho = −0.48, $S = 6,632$, $P = 0.01$), suggesting that the burrows are shallower, but not shorter, when constructed in compact, silty soil[11] (Fig. 1d). Overall, natural burrow shape and length are remarkably conserved in wild *P. polionotus*, yet variation in soil composition affects burrow depth. We therefore focused our genetic analyses on burrow length.

To measure the genetic component of complex burrowing, we first assayed the burrowing behaviour of *P. polionotus* and *P. maniculatus* in a single soil type under controlled laboratory conditions (see Methods). Briefly, each assay involved placing a mouse in a large, sand-filled enclosure for 46 h (two full-night activity periods). Then, we removed the mouse and made a polyurethane cast of the burrow (Supplementary Fig. 1c, d), which we then measured (Fig. 1b). We tested each animal in three consecutive assays. Captive-reared mice, which had never been exposed to sand or allowed to burrow, recapitulated their natural burrowing behaviour in our assays. We found no significant effects of assay number, sex, age, mass or enclosure on total burrow length (data not shown). Statistical comparisons of *P. polionotus* and *P. maniculatus* burrows revealed significant differences in total burrow length (Welch's two-tailed *t*-test, $t = 3.24$, df = 20.98, $P < 0.01$), consistent with previous results[15,16]. Moreover, total burrow length is composed of two main parts: entrance-tunnel length, which differed significantly between species (Fig. 2b; Welch's two-tailed *t*-test, $t = 6.72$, df = 24.39, $P < 5 \times 10^{-3}$), and escape-tunnel length—only *P. polionotus* constructed escape tunnels (Fig. 2c). Thus, despite having no previous experience with either sand substrate or our enclosures, these species built consistently distinct burrows in the laboratory.

To determine the inheritance patterns of this behavioural variation, we characterized the genetic architecture of burrowing differences observed between *P. polionotus* and *P. maniculatus*. First, we crossed the two species and assayed the burrows of their hybrid offspring (Supplementary Methods and Fig. 2a). First-generation (F$_1$) hybrids built entrance tunnels that were significantly longer (after Bonferroni correction, $\alpha = 1.67 \times 10^{-2}$) than *P. maniculatus* (Welch's two-tailed *t*-test, $t = 5.34$, df = 20.7, $P = 2 \times 10^{-4}$). But F$_1$ entrance lengths did not differ from the *P. polionotus* parent (one sample *t*-test, $\mu = 14$ cm, $t = -0.14$, df = 12, $P = 0.89$) or from the *P. polionotus* population mean (Fig. 2b; Welch's two-tailed *t*-test, $t = 2.35$, df = 22.7, $P = 0.03$), and all F$_1$ animals constructed escape tunnels (Fig. 2c). Thus, the alleles contributing to burrow size and shape seem to segregate in a dominant fashion.

We next crossed F$_1$ animals with *P. maniculatus* to create a recombinant backcross (BC) generation ($n = 272$). BC mice constructed entrance tunnels that varied continuously in length between the parental extremes, but approximately one of eight ($n = 36$) of the BC mice built *P. polionotus*-length tunnels (Fig. 2b; >14 cm in length), suggesting that only a few loci are necessary to generate this behaviour. In contrast, half (46%) of the BC mice built escape tunnels (Fig. 2c). This inheritance pattern is consistent with the action of either a single major-effect locus or of multiple loci that interact to create a threshold effect[19], such that only some loci need to be co-inherited to cause the expression of a trait. Finally, tunnel lengths and the presence of escape tunnels (that is, tunnel number) are only weakly correlated in BC mice
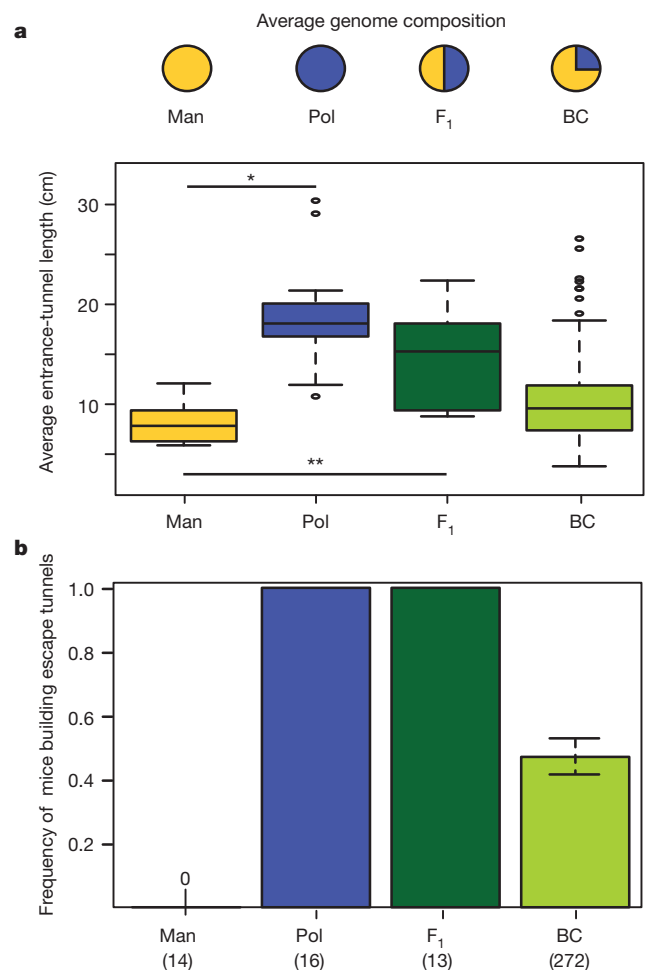

Figure 2 | Burrow variation across generations. a, Burrow dimensions of *P. maniculatus* (Man; yellow), *P. polionotus* (Pol; blue), F$_1$ hybrids (dark green) and progeny resulting from F$_1$ × *P. maniculatus* backcross (BC; light green). Pie charts depict average genome composition in each generation. Distributions of entrance-tunnel length (average of three trials for each individual tested) in the parental species, F$_1$ hybrids and BC animals are shown. Boxes represent interquartile ranges (median ± 2 s.d.). Significant *t*-tests, *$P = 5 \times 10^{-3}$, ** $P = 2 \times 10^{-4}$. **b**, The frequency of escape-tunnel construction is shown for the same individuals. Error bars represent mean ± standard error of the mean (s.e.m.) Sample sizes are listed in parentheses below.

(Supplementary Table 2). Therefore, the complex burrows of *P. polionotus* comprise at least two separate behavioural modules, one for tunnel length and one for the presence of an escape tunnel.

To identify the chromosomal locations involved and quantify the phenotypic effects of loci that influence these behavioural modules, we used a quantitative trait locus (QTL) mapping approach. We first identified 526 diagnostic single nucleotide polymorphisms (SNPs) using a double-digest restriction-site-associated-DNA (ddRAD) method[20], and genotyped the 272 BC mice for which we measured burrowing behaviour (that is, entrance-tunnel length and the presence of an escape tunnel; see Methods for details). Consistent with simple inheritance patterns predicted by burrow phenotypes in hybrid animals, we identified three genomic regions that contributed to variation in entrance-tunnel length, and a single region associated with escape-tunnel construction (Fig. 3a and Supplementary Fig. 5). All four QTLs are unlinked and segregate on separate chromosomes. The entrance-tunnel-length QTLs have similar effect sizes, interact additively (as determined by both a lack of epistasis between loci and graphical comparisons; Fig. 3b and Supplementary Methods), and together explain ~15% of tunnel-length variation (Table 1). Because approximately 24% of the variation in burrow length is likely to have a
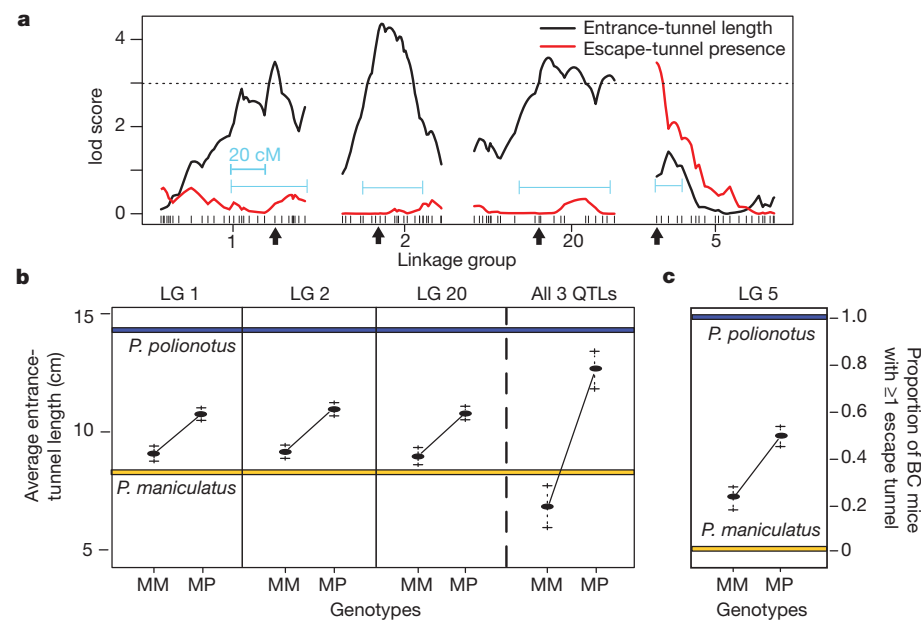
Figure 3 | QTL analysis of burrow variation. a, Linkage groups (LGs) 1, 2 and 20 harbour QTLs associated with log-transformed average entrance-tunnel length (black line). Linkage group 5 contains a single QTL associated with escape-tunnel presence (red line). Dotted line represents log odds ratio (lod) significance threshold (genome-wide $\alpha = 0.05$, lod = ~3.0). 1.5-lod confidence intervals and scale in centimorgans (cM) are shown in light blue. Dashes indicated genetic markers, and black arrows indicate markers used to define each QTL peak (used in **b**). **b**, Phenotypic effect of individual and combined QTLs (linkage groups 1, 2 and 20) on entrance-tunnel length in 272 BC mice. **c**, Proportion of BC animals that construct escape tunnels for each of the two genotypes. All error bars represent mean ± s.e.m. Blue and yellow lines represent average phenotype of the parents (pure species) used to found the cross. Genotypes are either homozygous *P. maniculatus* (MM) or heterozygous *P. maniculatus/polionotus* (MP).

genetic basis (as measured by the repeatability of burrow length in *P. polionotus*)[17], the QTLs that we report may explain more than half of the genetic variation for this trait (but see ref. 21). Even more notable, each *P. polionotus* allele increases entrance-tunnel length by, on average, 3 cm (Fig. 3b). Furthermore, BC individuals with all three *P. maniculatus* alleles versus those with one *P. polionotus* allele at each locus (that is, heterozygotes) nearly recapitulate the difference in entrance-tunnel length between the two species (that is, the mean phenotype of the parents used in the cross). Although it is possible that each QTL may harbour multiple, possibly interacting, mutations, these results are consistent with only a few genes (as few as three) being needed to explain the continuous tunnel-length variation observed in BC mice.

By contrast, QTL analysis identified only a single locus for the presence of escape tunnels on linkage group 5. Although this QTL explains only 6% of the escape-tunnel variance, BC individuals that inherited one *P. polionotus* allele at this locus were 30% more likely on average to build an escape tunnel than individuals with only *P. maniculatus* alleles (Fig. 3c). Given that the QTL is located at the end of a linkage group, it is possible that the causative mutation(s) is still far from our nearest marker. If this is the case, then we will have underestimated the phenotypic effect of this QTL, and it alone could explain the presence/absence of escape tunnels. An alternative explanation is that escape-tunnel construction is a threshold trait, but only a high-powered experiment would detect additional loci with either additive or epistatic effects on this behaviour. In either case, any additional QTLs are likely to have small phenotypic effects.

Our results show that QTLs are specific to either entrance-tunnel length or escape-tunnel presence, highlighting the modular nature of the complex *P. polionotus* burrowing behaviour. Moreover, all alleles derived from the *P. polionotus* parent cause an increase in trait value (that is, longer entrance tunnels or the addition of an escape tunnel; Fig. 3b), consistent with the role of natural selection in driving the evolution of the complex burrow architecture. Together, these results show that two aspects of burrow architecture evolved independently, as they are genetically discrete, suggesting that entrance-tunnel length and the presence of an escape tunnel are both ecologically important components of the complex *P. polionotus* burrows.

Although progress has been made towards understanding the genetic underpinnings of innate behaviours, most studies have been restricted to a few laboratory-based model systems[22] or candidate gene approaches in more diverse taxa[23]. By taking advantage of variation in an extended phenotype, in this case burrow architecture in *Peromyscus*, we have gained insight into how a seemingly complex behaviour evolves in the wild. First, we show that the complex, derived burrowing behaviour of *P. polionotus* largely results from evolution at a surprisingly small number of loci. Second, we demonstrate that burrowing behaviour is composed of distinct behavioural modules, which are controlled by independent genetic loci. This is consistent with a century-old observation that elaborate animal behaviour, such as the courtship rituals of grebes, may evolve by integrating simple behaviours[24]. Although genetic modularity has been shown to be important for the evolution of complex morphological traits[25] and laboratory-based behaviours[26], our results extend this pattern to wild

Table 1 | Models for individual and combined QTL effects

Trait	Chromosome	Position	Genotype completeness (%)	lod score	PVE	Additive effect
Log$_{10}$ (average entrance-tunnel length)	1	75.9	77.6	2.48	3.6	0.063
	2	25.1	73.9	3.34	4.9	0.072
	20	59.6	58.8	2.47	3.6	0.064
Full entrance-tunnel-length model	-	-	-	9.55	14.8	-
Escape-tunnel presence (binary)	5	0	60.3	3.38	6.2	0.943

Full model includes all three QTLs. lod, log odds ratio; PVE, percentage of phenotypic variance explained.

behaviours. Thus, we suggest that the behavioural diversity observed in nature may often evolve by accumulating and combining alleles, each with modular effects.

METHODS SUMMARY

Experimental design. We originally obtained outbred stocks of *Peromyscus* from the *Peromyscus* Genetic Stock Center. To start the genetic cross, we mated a single female *P. maniculatus bairdii* with a male *P. polionotus subgriseus* to generate 13 F_1 hybrids, which were then backcrossed to *P. maniculatus* to produce 272 BC-generation hybrids. We assayed burrowing behaviour following our previously described methods[16]. In brief, we placed a single, virgin mouse (60–90 days of age) in an $1.2 \times 1.5 \times 1.1$ m enclosure filled with food, water, nesting material and 700 kg of hard-packed sand under constant temperature and light cycle. Each trial lasted approximately 46 h, and at the end of each trial, we made a polyurethane cast of the resulting burrow[16,27]. From each cast, we measured the length of the entrance tunnel (from the surface to the nest chamber; Fig. 1b) and recorded the presence/absence of an escape tunnel. All mice were assayed in three, consecutive 2-day trials, each time in a new, randomly assigned enclosure.

Genotyping. We extracted DNA using a high-throughput automated phenol-chloroform method (Autogen). We genotyped all individuals using a ddRAD-sequencing approach[20], which identified 526 SNPs with fixed nucleotide differences between the parental species. We used the genotypes of hybrids to estimate genetic linkage among markers. This produced a map containing 24 linkage groups, corresponding to the 24 chromosomes in *P. maniculatus* and *P. polionotus*, with a total map length of 1835.5 cM (Supplementary Fig. 2).

Data analysis. Burrow lengths were log-normalized before correlation and QTL analyses. We performed all statistical tests in the R statistical package[28]. QTLs associated with burrow phenotypes were identified using r/QTL[29]. Specifically, we sequentially performed Haley–Knot regressions, interval mapping, and interval mapping with imputation of missing genotypes, and report QTLs that are consistent across all three methods. We used permutation tests to determine genome-wide statistical significance for each QTL.

Full Methods and any associated references are available in the online version of the paper.

1. Boake, C. R. B. *et al.* Genetic tools for studying adaptation and the evolution of behavior. *Am. Nat.* **160,** S143–S159 (2002).
2. West-Eberhard, M. J. *Developmental Plasticity and Evolution* (Oxford Univ. Press, 2003).
3. Mackay, T. F. C. The genetic architecture of quantitative traits. *Annu. Rev. Genet.* **35,** 303–339 (2001).
4. Dawkins, R. *The Extended Phenotype* (W. H. Freeman, 1982).
5. Hansell, M. H. *Animal Architecture* (Oxford Univ. Press, 2005).
6. Lorenz, K. Z. The evolution of behaviour. *Sci. Am.* **199,** 67–78 (1958).
7. Peichel, C. L. *et al.* The genetic architecture of divergence between threespine stickleback species. *Nature* **414,** 901–905 (2001).
8. Steiner, C. C., Weber, J. N. & Hoekstra, H. E. Adaptive variation in beach mice caused by two interacting pigmentation genes. *PLoS Biol.* **5,** 1880–1889 (2007).
9. Sumner, F. B. & Karol, J. J. Notes on the burrowing habits of *Peromyscus polionotus.* *J. Mamm.* **10,** 213–215 (1929).
10. Hayne, D. W. Burrowing habits of *Peromyscus polionotus.* *J. Mamm.* **17,** 420–421 (1936).
11. Rand, A. L. & Host, P. Mammal notes from Highland County, Florida. *Bull. Am. Mus. Nat. Hist.* **80,** 1–21 (1942).
12. Houtcooper, W. C. Rodent seed supply and burrows of *Peromyscus* in cultivated fields. *Proc. Indiana Acad. Sci.* **81,** 348–389 (1971).
13. Schwartz, C. W. & Schwartz, E. R. *The Wild Mammals of Missouri* (Univ. Missouri Press, 1981).
14. Baker, R. H. in *Biology of Peromyscus (Rodentia)* (ed. King, J. A.) (American Society of Mammalogists, 1968).
15. Wolfe, J. L. & Esher, R. J. Burrowing behaviour of old-field mice (*Peromyscus polionotus*): a laboratory investigation. *Bio. Behav.* **2,** 343–351 (1977).
16. Dawson, W. D., Lake, C. E. & Schumpert, S. S. Inheritance of burrow building in *Peromyscus. Behav. Genet.* **18,** 371–382 (1988).
17. Weber, J. N. & Hoekstra, H. E. The evolution of burrowing behavior in deer mice. *Anim. Behav.* **77,** 603–609 (2009).
18. Tan, K. H. *Soil Sampling, Preparation, and Analysis* (CRC Press, 2005).
19. Wright, S. A mutation of the guinea pig, tending to restore the pentadactyl foot when heterozygous, producing a monstrosity when homozygous. *Genetics* **20,** 84–107 (1935).
20. Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS ONE* **7,** e37135 (2012).
21. Beavis, W. D. in *Molecular Dissection of Complex Traits* (ed. Paterson, A. H.) 431–528 (CRC Press, 1998).
22. Bendesky, A. & Bargmann, C. I. Genetic contributions to behaviour at the gene–environment interface. *Nature Rev. Genet.* **12,** 809–820 (2011).
23. Fitzpatrick, M. J. *et al.* Candidate genes for behavioural ecology. *Trends Ecol. Evol.* **20,** 96–104 (2005).
24. Huxley, J. S. The courtship habits of the great crested grebe (*Podiceps cristatus*) with an addition to the theory of sexual selection. *Proc. Zool. Soc. Lond.* **35,** 253–291 (1914).
25. Mallarino, R. *et al.* Two developmental modules establish 3D beak shape variation in Darwin's finches. *Proc. Natl Acad. Sci. USA* **108,** 4057–4062 (2011).
26. Xu, X. *et al.* Modular genetic control of sexually dimorphic behaviors. *Cell* **148,** 596–607 (2012).
27. Felthauser, M. & McInroy, D. Mapping pocket gopher burrow systems with expanding polyurethane foam. *J. Wildl. Manage.* **47,** 555–558 (1983).
28. R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2011).
29. Broman, K. W., Wu, H., Sen, Ś. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19,** 889–890 (2003).

Supplementary Information is available in the online version of the paper.

Author Contributions J.N.W. and H.E.H. conceived and designed the experiments. B.K.P. and J.N.W. generated the ddRAD genotypes. J.N.W. performed the behaviour experiments and analysed the genetic and behavioural data. J.N.W. and H.E.H. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.E.H. (hoekstra@oeb.harvard.edu).

METHODS

Field observations and soil analysis. We cast and excavated *P. polionotus* burrows at eight sites, distributed broadly across the species range. First, we removed mice from active burrows by inserting plastic tubing into a burrow entrance tunnel while slowly excavating the tunnel (see Supplementary Video for the method conducted in a laboratory environment). The mice generally exited burrows through their escape tunnels when our excavation neared the nest chamber. Once the mice were removed, we filled the remaining burrow cavity with expansive polyurethane filling foam (either from Hilti, or Great Stuff, Dow Chemicals)[27]. After the foam hardened, we excavated the cast and measured the entrance tunnel, the total burrow length, as well as the depth of the lowest point in the nest from the surface (Supplementary Table 1).

In addition to burrow measurements, we collected and analysed samples of undisturbed soil that immediately surrounded the nest chambers of excavated burrows. Before analysing the soil, we sifted each sample through a standardized sieve (2 mm diameter) and then oven dried the remaining soil. We then performed a gravimetric particle analysis on the dried soils (using a hydrometer method[18]) to quantify clay, sand and silt content (Supplementary Table 1).

Statistical analyses of soil and burrow variation. Initially, we $\log_{10}$-transformed all continuous variables to make them normally distributed. For the soil and natural burrow measurements, we used conservative, non-parametric tests to analyse these data because we had small sample sizes. Specifically, we used Spearman tests for analyses of correlation and Kruskal–Wallis tests for differences in variation among groups. For the data collected in the laboratory, we calculated Pearson correlation coefficients among all behavioural variables, and also tested for correlations between burrow measures and soil moisture, animal mass and animal age. We found no significant correlations between total burrow length and moisture, mass or age ($P > 0.05$). Similarly, an analysis of variance (ANOVA) showed no differences in the total length of burrows constructed in different enclosures ($P > 0.05$). We performed all statistical tests using R[28].

Genetic cross. We performed experiments using outbred mice originally obtained from the *Peromyscus* Genetic Stock Center (PGSC). We initially crossed a single female *P. maniculatus bairdii* with a male *P. polionotus subgriseus* to generate 13 first-generation (F_1) hybrids, which we then backcrossed to *P. maniculatus* to produce 272 BC-generation hybrids. We followed standard husbandry protocols used by the PGSC, and Harvard University's Institutional Animal Care and Use Committee approved our experimental methods (protocol 27-09).

Burrowing assays. We housed animals and assayed their burrowing behaviour under controlled laboratory conditions. Specifically, we set the light cycle to 16 h of light and 8 h dark, and room temperature and humidity remained constant at 21–24 °C and 30–37%, respectively. When mice reached an adult age of 60–90 days, we tested their burrowing behaviour in a semi-natural enclosure (described later). Each burrowing trial involved introducing a single mouse to an enclosure at the start of a dark cycle, and then removing the animal approximately 46 h later. After removing the mice, we made casts of burrows by injecting them with polyurethane filling foam (Hilti). We measured the entrance tunnel and total tunnel length of each burrow directly from the casts, and also recorded the presence or absence of an escape tunnel. We conducted three consecutive burrowing trials on each animal, and moved mice to a different enclosure for each trial.

We assayed burrowing behaviour in ten, $1.2 \times 1.5 \times 1.1$ m enclosures, each filled with 700 kg of Premium Playground Sand (Pharmaserv). Before introducing the mice, we contoured the soil surface into three equally sized sections: a flat lower surface; a ~50-cm-high hill at an angle of 60° to horizontal; and a flat upper surface (Fig. 1a). Each enclosure included approximately 5 g of standard rodent food, a 5 cm² cotton nestlet (PharmaServ) and a water bottle. At the completion of each trial, we removed all debris from the enclosures, measured soil moisture with a Hydrosense probe (Campbell Scientific), and then thoroughly mixed the sand. We continuously added sand to the enclosures throughout the experiment to maintain constant soil and moisture levels (5–9% and 17–22% water content on the upper and lower surfaces, respectively).

Genotyping. We initially identified 1,700 SNPs, each fixed within but different between *P. maniculatus* and *P. polionotus*, using a modified RAD-tag approach[20].

Briefly, for each individual, we digested ~1 µg of genomic DNA with two restriction endonucleases (100 units of EcoR1-HF and 20 units of Msp1). We ligated the resulting fragments to sequencing adapters containing unique barcodes for each sample. We then pooled ligation products among individuals and isolated fragments in the size range of 280–320 bp using a Pippin Prep electrophoresis platform (Sage Biosciences). Finally, we amplified the remaining fragments using a Phusion High Fidelity PCR Kit (ThermoFisher Scientific) and sequenced the resulting libraries on a Genome Analyzer II (Illumina). For a more complete description of the computational methods used to analyse short-read sequences and to determine genotypes, see ref. 20.

Linkage map construction. We constructed a linkage map in R/qtl[29] using genotypes from 1,700 markers scored in BC-generation individuals. Initially, we calculated the fraction of recombination events and lod score between all marker pairs. We identified 97 markers with genotype information identical to another marker and removed them from the map. We then grouped markers by varying the maximum fraction of recombination events and minimum lod score allowed between markers on the same linkage group. Because the karyotypes of both species are known ($n = 24$ chromosomes), we varied recombination parameters until we recovered a map with 24 linkage groups, each comprising at least 30 markers. This map also contained linkage groups with fewer than 10 markers; we removed the markers in small linkage groups. We ordered the remaining markers by individually scanning linkage groups in overlapping windows of 4–8 markers and minimizing the frequency of recombination events between markers in each window. Next, we removed all markers genotyped in fewer than 100 individuals (Supplementary Fig. 2) as well as all markers with high error rates[30]. Finally, we imputed missing genotypes among all markers that shared identical map locations and then pruned the remaining marker set to reflect an average intermarker distance of ~5–10 cM. The final linkage map contained 526 markers with an average of 390 (s.d. ± 97) genotypes per individual (Supplementary Fig. 3). Finally, we evaluated our genotyping error rate by comparing the likelihood of our marker data, given our estimated linkage map, under different error regimes.

QTL mapping. We performed Haley–Knott regressions and interval mapping analyses sequentially in R/qtl[29] to identify QTLs contributing to burrowing differences. Using permutation tests, we determined the genome-wide significance level for association between markers and phenotypes ($\alpha = 0.05$, $n = 2,000$ permutations) as lod = 3.05 and 3.08 for average entrance-tunnel length and binary escape-tunnel presence, respectively[31]. For mapping of escape-tunnel presence, we excluded 29 animals that had unclear escape-tunnel phenotypes (criteria: built an escape tunnel in only one out of three trials and the escape tunnel was ≤4 cm long). lod scores across all linkage groups for both entrance-tunnel length and escape-tunnel presence are shown in Supplementary Fig. 4. With 272 animals in a backcross design, we had 80% power to detect QTLs with effect sizes of >0.3 s.d., assuming the trait is ~70% heritable[32]. We also scanned for pairwise interactions between loci across all linkage groups (using R/qtl[29]), as well as for QTLs that segregated differentially among our BC families due to the outbred nature of our parental stocks (using QTLrel[33,34]). We found no significant evidence for epistasis or for family differences in QTLs. To determine the additive effect of having different genotypes at markers under QTL peaks, we used the fitqtl() and refineqtl() functions in R/qtl[29] to construct models that estimated the percentage of phenotypic variance explained by each QTL (Table 1).

30. Lincoln, S. E. & Lander, E. S. Systematic detection of errors in genetic linkage data. *Genomics* **14,** 604–610 (1992).
31. Doerge, R. W. & Rebai, A. Significance thresholds for QTL interval mapping tests. *Heredity* **76,** 459–464 (1996).
32. Sen, Ś., Satagopan, J., Broman, K. W. & Churchill, G. A. *R/qtlDesign: Inbred Line Cross Experimental Design* (UC San Francisco: Center for Bioinformatics and Molecular Biostatistics, 2006).
33. Cheng, R. *et al.* Genome-wide association studies and the problem of relatedness among advanced intercross lines and other highly recombinant populations. *Genetics* **185,** 1033–1044 (2010).
34. Cheng, R., Abney, M., Palmer, A. A. & Skol, A. D. QTLRel: an R package for genome-wide association studies in which relatedness is a concern. *BMC Genet.* **12,** 66 (2011).

CAREERS

EDUCATION

Free-range learning

The budding field of informal science education offers varied research paths but uncertain funding.

BY VIRGINIA GEWIN

Zombies are on the attack. Victims' organs must be replenished to save humanity from the undead. But only players with an understanding of cells, tissues and organs — even the latest techniques to induce pluripotent stem cells — can regenerate the lost body parts.

This is the challenge set forth in *Progenitor X*, a problem-solving video game in which players learn how to reprogram more than 200 human cell types. "People play *Progenitor X* because it is fun to learn some things about science as they play," says game developer Kurt Squire, who works at the University of Wisconsin–Madison.

Squire is one of a growing number of informal-science-education (ISE) researchers, who focus on providing learning experiences outside of the formal classroom environment. The field encompasses a diverse group of talents, including video-game developers, as well as people who design, develop and evaluate science-education experiences for museums, television shows, after-school programmes, science festivals and national parks. "The learning that takes place in museums and science centres is the same learning as in school — brains generally work the same no matter the location," says Justin Dillon, an ISE researcher at Kings College London, which is launching a bachelor's degree in science engagement and communication in

September. "But those experiences can reinforce or deepen the learning."

ISE is increasingly gaining traction: both government and non-profit funders are supporting activities; some museums are hiring more research staff; and the number of training opportunities is increasing. Attempting to organize the field is the Washington DC-based Center for Advancement of Informal Science Education (CAISE), which was formed in 2007 with support from the US National Science Foundation (NSF). CAISE aims to raise the field's profile by bringing together diverse sectors under the ISE umbrella to garner more funding, formalize training opportunities and establish methods to assess the success of various outreach projects.

However, in the face of strained science budgets, this budding field has been prone to funding ups and downs. Government agencies might perceive their activities as lower priorities. Practitioners have increasingly forged partnerships with industry or private foundations that are eager to cultivate a science-savvy citizenry and workforce. Although diverse job opportunities exist, seizing them requires connections, flexibility and creativity.

ACADEMIC RESEARCH

The diversity of career tracks may be the field's biggest perk. Science centres, zoos, aquaria, video-game companies, television shows and non-profit organizations are all looking to develop new materials to keep visitors and users engaged. Research-oriented jobs are among the fastest-growing areas of employment. "Funders want people to fundamentally look at the learning question — it's not just about building an exhibit any more to get NSF funding," says Rhiannon Crain, who trained with Dillon and is now leading a citizen-science project called Yard Map. Run from the Cornell Lab of Ornithology in Ithaca, New York, Yard Map asks users to provide data about the effect of landscape changes on wildlife, and teaches them how best to create a bird habitat.

ISE practitioners might quantify, for example, how access to scientific instruments or engaging in hands-on exploration — through laboratory experiments or forensic-investigation simulations — can spark curiosity and strengthen someone's comprehension of science. Kevin Crowley, an education researcher at the University of Pittsburgh in Pennsylvania, has partnered with the Carnegie Museum of Natural History in Pittsburgh to help the ▶

museum to enrich civic debates that involve scientific evidence. For example, to increase climate-change literacy in urban communities, the museum is breaking down the huge, complex topic into small pieces, focusing on how climate change will affect Pittsburgh itself.

Approaches to ISE, and therefore to training options, differ. A group led by Lynn Dierking at Oregon State University in Corvallis focuses on what motivates people to seek learning experiences in daily life, a concept she calls "free-choice learning". Dierking's group studies how to create opportunities that pique human curiosity, such as during visits to state parks.

The university offers a six-course certificate programme, an online master's degree and a formal PhD. Trainees can have a science or social-science background. The online master's is designed to provide the skills to enhance visitor learning at museums, national parks or nature centres. PhD candidates typically conduct research at an ISE outlet (such as a museum or an after-school programme), and recipients generally go into academia or become evaluators.

But the academic track often offers little job security. "Unfortunately, most of those are not tenure-track positions," says Crain. Those who do find university-based work are typically adjunct faculty who rely heavily year-to-year on grant money. Squire's students end up not only in academia but also in the entertainment industry or starting their own companies.

OUTSIDE ACADEMIA

Museums, meanwhile, particularly larger ones, are building up internal research and evaluation teams to track, for example, visitor experiences through surveys. The Science Museum of Minnesota in St Paul and the Exploratorium in San Francisco, California, have teams of 12 to 14 who use this visitor feedback to hone exhibits. Larger museums have programme or exhibit departments — many of which are staffed by scientists with doctoral degrees — that develop and design exhibits, a process that can take two to three years from conception to completion. Smaller museums often have to hire external evaluators or companies to design and fabricate their exhibits. (Although some museums, notably The Field Museum in Chicago, Illinois, are facing harsh budget realities that have meant cuts affecting researchers and other employees (see *Nature* http://doi.org/j6q; 2012).)

The increasing demand for consultant evaluators is a bright spot that seems poised to continue. "The need for evaluators in ISE is growing even when, and perhaps because of, decreases in funding," says Kirsten Ellenbogen, senior director of lifelong learning at the Science Museum of Minnesota. It's more urgent to demonstrate impacts than in past years, she adds. Evaluators collect data about the learning experiences of museum visitors to show whether the programme or exhibit has effectively educated or enlightened. Evaluation is often crucial to satisfying grant requirements. "In the current funding environment, accountability is critical," agrees Ellen McCallie, programme director with NSF's division of research on learning in Arlington, Virginia.

As a result, evaluators can find work as project-based consultants or join evaluation

> *"We are really interested in understanding what motivates and sustains interest in science."*

firms that are popping up. The research institute SRI International, headquartered in Menlo Park, California, continues to grow — hiring people with social-science skills and a passion for communicating science. A few universities, notably Boston College in Massachusetts, the University of California, Berkeley, and Claremont Graduate University in California, offer PhD-level training that specializes in designing quantitative-research studies to analyse visitor data.

FUNDING FLUCTUATIONS

Regardless of the track, funding is a challenge. The NSF cut its funding request for the Advancing Informal STEM (Science, Technology, Engineering and Maths) Learning programme for the fiscal year 2013 by 22% from 2012 numbers, to US$43 million. Funding is equally uncertain outside the United States. It is not yet clear whether 'Science in Society' — one of the two main ISE funding schemes under the European Framework Programme — will continue to exist in 2014.

Yet, at the same time, funders such as the NSF broadly encourage the use of ISE outreach and evaluation techniques across other NSF programmes and centres. For example, the NSF's Centers for Chemical Innovation programme requires a robust effort in informal science communication or ISE. The programme's director, Katharine Covert, says that there is a supplemental funding opportunity of up to $150,000 for the centres to develop collaborations with an ISE organization. And a new NSF-funded initiative — Science Education For New Civic Engagements and Responsibilities (SENCER) — will establish partnerships between higher education and ISE organizations to find innovative ways to share research and interact with the public. The next round of SENCER funding, as part of a "Civic Engagement Partnership" programme, will provide six $50,000 partnership awards (applications are due by 15 March) to ISE–higher-education collaborations.

"These days, informal-science educators have to be flexible and imaginative about how to obtain funding in this research area; ideally, they are able to piggyback onto existing activities," says Karen Bultitude, director of research in the department of science and technology studies at University College London. She expects that the UK's new Research Excellence Framework, which requires researchers to demonstrate the impact of their research outside academia, will encourage other groups to collaborate with ISE outreach projects. "Almost all of academia in the UK is desperately interested in demonstrating impacts of their work on wider groups," says Bultitude.

In some cases, private funders have stepped in. The Wellcome Trust in London has spent upwards of £50 million (US$80 million) over the past decade on schemes that include media fellowships for scientists and, most recently,

The zombie video game *Progenitor X* helps players to learn how cell types are formed.

WISCONSIN INST. DISCOVERY

engagement fellowships to support senior researchers who are developing a programme or projects to share their expertise with the public. One recipient, Roger Kneebone, based at Imperial College London, is a surgeon-turned-public-engagement fellow who plans to use live surgical simulations to engage the public. Clare Matterson, director of medical humanities and engagement at the Wellcome Trust, hopes that two reports published by the trust in November 2012 — *Analysing the UK Science Education Community: the Contribution of Informal Providers* and *Review of Informal Science Learning* — will prompt more grant submissions in the United Kingdom, and for ISE in general, of which there is a dearth at the moment, she says. "What is needed is a greater number of high-quality research-based proposals so we have a better understanding of how children learn through informal science experiences," she adds. "These don't have to be education groups; they could be people working in psychology or neuroscience, studying how people respond to different types of learning."

In the United States, the Gordon and Betty Moore Foundation and the Noyce Foundation, both in Palo Alto, California, are funding ISE efforts ranging from citizen-science engagement to helping teachers in training learn how to develop their own interactive science-lesson plans. "We are really interested in understanding what motivates and sustains interest in science," says Janet Coffey, a programme officer at the Gordon and Betty Moore Foundation.

Industry is also experimenting with ISE in an effort to reach the public. Novo Nordisk, headquartered in Bagsvaerd, Denmark, for example, finances the Steno Health Promotion Centre in Gentofte, Denmark, which runs a 33-million-krone (US$5.7 million) project called PULSE at the Experimentarium in Copenhagen, to promote healthy lifestyles for families. The project will engage and track families with children aged 6–12 years old from a variety of socioeconomic backgrounds as they design and develop healthful eating and exercise strategies. The museum is increasingly "looked at as a serious partner instead of simply an institute for kids," says Sheena Laursen, director of international projects at the Experimentarium.

Many hope that continued research efforts that document how ISE increases scientific literacy or encourages students to pursue STEM careers will help the field to grow. "Anyone who cares about scientific literacy and STEM workforce development," says Ellenbogen, "should pay attention to ISE." ∎

Virginia Gewin *is a freelance writer in Portland, Oregon.*

Q&A Francis Collins
Changes at the NIH

For years, the US National Institutes of Health has struggled with promoting non-academic career tracks for biomedical scientists, gauging the supply of PhD holders and demand for research jobs, enticing under-represented minorities into science and establishing funding avenues for early-career researchers. Hoping to bring some evidence-based clarity to these issues, NIH director Francis Collins asked two working groups of the NIH Advisory Committee to study the issues and make recommendations. They released their recommendations in two reports in June; Collins responded in December. The NIH has decided to take measures that include raising its postdoc stipend, increasing the number of grants that encourage early-career independence and offering 25 institutional grants, each worth about US$250,000, to support training programmes that prepare students for a broad range of research-related careers, including non-academic paths.

What did you learn from the advisory reports?
The working groups put in a lot of effort collecting data to figure out what career paths PhDs were ultimately heading down. Only about 23% of US-trained biomedical PhD holders were in academic tenure or tenure-track positions in 2008, they found. Many end up in research-related positions, in industry, government, teaching, science policy, science journalism and other science-related professions. Training programmes have tended to view those tracks as secondary and have even sent messages that they are second-rate. So we need to retool to expose trainees to multiple pathways, rather than simply producing clones of their principal investigators.

What might the $250,000 institutional grants involve?
We propose to have institutions provide actual exposure to careers outside a university setting, to help them to think more creatively about the skills and knowledge that PhD holders might need for the job market. For example, students and postdocs may spend time at a pharmaceutical company, develop entrepreneurial skills or learn aspects of science-policy development.

Where will you get the money to support these programmes?
We can do it with the NIH Common Fund, which is dedicated to innovative initiatives, especially for topics that are relevant to multiple NIH institutes or centres. We will have to balance that with the need to continue other programmes through the Common Fund.

Unemployment is low among scientists in the United States, at 4–5%. Is the priority, then, to change expectations among those considering the academic track?
Yes, that is the intention. Unemployment is actually about 2% among biomedical PhD ▶

engagement fellowships to support senior researchers who are developing a programme or projects to share their expertise with the public. One recipient, Roger Kneebone, based at Imperial College London, is a surgeon-turned-public-engagement fellow who plans to use live surgical simulations to engage the public. Clare Matterson, director of medical humanities and engagement at the Wellcome Trust, hopes that two reports published by the trust in November 2012 — *Analysing the UK Science Education Community: the Contribution of Informal Providers* and *Review of Informal Science Learning* — will prompt more grant submissions in the United Kingdom, and for ISE in general, of which there is a dearth at the moment, she says. "What is needed is a greater number of high-quality research-based proposals so we have a better understanding of how children learn through informal science experiences," she adds. "These don't have to be education groups; they could be people working in psychology or neuroscience, studying how people respond to different types of learning."

In the United States, the Gordon and Betty Moore Foundation and the Noyce Foundation, both in Palo Alto, California, are funding ISE efforts ranging from citizen-science engagement to helping teachers in training learn how to develop their own interactive science-lesson plans. "We are really interested in understanding what motivates and sustains interest in science," says Janet Coffey, a programme officer at the Gordon and Betty Moore Foundation.

Industry is also experimenting with ISE in an effort to reach the public. Novo Nordisk, headquartered in Bagsvaerd, Denmark, for example, finances the Steno Health Promotion Centre in Gentofte, Denmark, which runs a 33-million-krone (US$5.7 million) project called PULSE at the Experimentarium in Copenhagen, to promote healthy lifestyles for families. The project will engage and track families with children aged 6–12 years old from a variety of socioeconomic backgrounds as they design and develop healthful eating and exercise strategies. The museum is increasingly "looked at as a serious partner instead of simply an institute for kids," says Sheena Laursen, director of international projects at the Experimentarium.

Many hope that continued research efforts that document how ISE increases scientific literacy or encourages students to pursue STEM careers will help the field to grow. "Anyone who cares about scientific literacy and STEM workforce development," says Ellenbogen, "should pay attention to ISE." ∎

Virginia Gewin *is a freelance writer in Portland, Oregon.*

Q&A Francis Collins
Changes at the NIH

For years, the US National Institutes of Health has struggled with promoting non-academic career tracks for biomedical scientists, gauging the supply of PhD holders and demand for research jobs, enticing under-represented minorities into science and establishing funding avenues for early-career researchers. Hoping to bring some evidence-based clarity to these issues, NIH director Francis Collins asked two working groups of the NIH Advisory Committee to study the issues and make recommendations. They released their recommendations in two reports in June; Collins responded in December. The NIH has decided to take measures that include raising its postdoc stipend, increasing the number of grants that encourage early-career independence and offering 25 institutional grants, each worth about US$250,000, to support training programmes that prepare students for a broad range of research-related careers, including non-academic paths.

What did you learn from the advisory reports?
The working groups put in a lot of effort collecting data to figure out what career paths PhDs were ultimately heading down. Only about 23% of US-trained biomedical PhD holders were in academic tenure or tenure-track positions in 2008, they found. Many end up in research-related positions, in industry, government, teaching, science policy, science journalism and other science-related professions. Training programmes have tended to view those tracks as secondary and have even sent messages that they are second-rate. So we need to retool to expose trainees to multiple pathways, rather than simply producing clones of their principal investigators.

What might the $250,000 institutional grants involve?
We propose to have institutions provide actual exposure to careers outside a university setting, to help them to think more creatively about the skills and knowledge that PhD holders might need for the job market. For example, students and postdocs may spend time at a pharmaceutical company, develop entrepreneurial skills or learn aspects of science-policy development.

Where will you get the money to support these programmes?
We can do it with the NIH Common Fund, which is dedicated to innovative initiatives, especially for topics that are relevant to multiple NIH institutes or centres. We will have to balance that with the need to continue other programmes through the Common Fund.

Unemployment is low among scientists in the United States, at 4–5%. Is the priority, then, to change expectations among those considering the academic track?
Yes, that is the intention. Unemployment is actually about 2% among biomedical PhD ▶

holders. But although the training is aimed almost entirely at preparing people for tenure-track positions in academia, only a minority of US-trained biomedical researchers end up there.

The working groups recommend supporting students and postdocs through trainee grants to shorten the time it takes to get a degree and improve the experience. Why have you declined to do so?

It is all about logistics. Something like 80% of postdocs and two-thirds of graduate students and trainees are associated with research grants — how do we put them on training grants instead? The shift would be an enormous administrative challenge. Instead, we are asking institutions to apply training-focused principles to all students and postdocs. For example, we recommend that all trainees have individual development plans to track how they are doing.

Do you think that it is the NIH's job to decide how many biomedical PhDs are needed?

This question fascinates and troubles me. It seems as though not a week goes by without someone saying that there are too many PhDs and someone else saying there are not enough. They can't both be right.

We are planning to establish a sophisticated economic-modelling unit at the NIH to determine supply and demand. We are missing a lot of data. We need to determine who is interested in going into PhD and medical doctorate–PhD programmes, and estimate the current and future demand for their skills in all sectors.

Does science need another established career track? What would it look like?

The working groups looked at career tracks for people who don't see themselves as the head of a lab. For example, on the NIH intramural campus in Bethesda, Maryland, we have a number of staff scientists. I have two in my own laboratory. I think the output per dollar spent for these scientists is very high. The working groups recommended that the peer-review process should recognize the contributions of staff scientists and consider them more favourably. At the NIH, we can encourage study sections to recognize the value of these scientists.

What is the 'diversity challenge' that the biomedical community is facing?

The diversity of the US biomedical research workforce is woefully short of that in the overall population. That is damaging our ability to carry out research as effectively as we would like. For example, there is evidence that individuals from under-represented groups tend to be interested in disparities in health between demographic groups, because they often come from communities that have experienced them. The absence of effective programmes

to recruit and retain such individuals leaves a gap in the biological sciences.

The NIH and other agencies have had various diversity initiatives, with limited effect. Can the NIH effect a sea change?

A sea change has been proposed by the diversity working group. On entrance to college, under-represented groups show a lot of interest in science. But we are losing a lot of talented people during the undergraduate years. Much of that is due to the lack of exposure to real research experience. Financial concerns about a long training period are also a factor.

We have proposed a programme called BUILD (Building Infrastructure Leading to Diversity), coming in the next 18 months. Undergraduate students at institutions that are not traditionally research-intensive will have an opportunity to do summer research in their chosen field and have their junior and senior years at college paid for. They will pay this back with two years of research in a full-time job.

The NIH is proposing to launch a mentoring network. What is its aim?

The idea is to give young scientists around the country a chance to be mentored by people who have similar interests and backgrounds. We will make awards to organizations to run the network after a competitive process.

According to a study published in 2011, African-Americans have lower rates of success with grant applications than other applicants. What is being done to address this?

We are planning more peer-review experiments, following up on concerns about possible unconscious bias in science [based on the findings in D. K. Gunther *et al. Science* **333**, 1015–1019 (2011)]. We at the NIH found the evidence alarming and we are determined to get to the bottom of it.

Do you discuss career paths with the postdocs in your lab?

We talk about career paths regularly, but I'm not sure that we cover all of the options effectively. This is perhaps a wake-up call for me. Most of my trainees end up in academia, but I probably haven't done a great job of exposing them to other options.

How would you advise a postdoc struggling with career options?

It is hard to generalize; you need to take into account their dreams, their talents and the contributions that they can make in different settings. I worry that a number of them are receiving the message that if they don't get a tenure-track position, they have failed. The good news is that nearly all postdocs are likely to be employed in interesting positions, but many will not travel a narrow academic path. ■

INTERVIEW BY GENE RUSSO